Measuring the effects of spectral smearing and enhancement on speech recognition in noise for adults and children

Susan Nittrouer,[a] Eric Tarr,[b] Taylor Wucinich, Aaron C. Moberly, and Joanna H. Lowenstein

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

(Received 22 August 2014; revised 27 February 2015; accepted 2 March 2015)

Broadened auditory filters associated with sensorineural hearing loss have clearly been shown to diminish speech recognition in noise for adults, but far less is known about potential effects for children. This study examined speech recognition in noise for adults and children using simulated auditory filters of different widths. Specifically, 5 groups (20 listeners each) of adults or children (5 and 7 yrs), were asked to recognize sentences in speech-shaped noise. Seven-year-olds listened at 0 dB signal-to-noise ratio (SNR) only; 5-yr-olds listened at +3 or 0 dB SNR; and adults listened at 0 or −3 dB SNR. Sentence materials were processed both to smear the speech spectrum (i.e., simulate broadened filters), and to enhance the spectrum (i.e., simulate narrowed filters). Results showed: (1) Spectral smearing diminished recognition for listeners of all ages; (2) spectral enhancement did not improve recognition, and in fact diminished it somewhat; and (3) interactions were observed between smearing and SNR, but only for adults. That interaction made age effects difficult to gauge. Nonetheless, it was concluded that efforts to diagnose the extent of broadening of auditory filters and to develop techniques to correct this condition could benefit patients with hearing loss, especially children.
[http://dx.doi.org/10.1121/1.4916203]

[EB] Pages: 2004–2014

I. INTRODUCTION

A continuing challenge for clinicians and educators concerned with helping children with hearing loss attain their optimal potentials is the fact that auditory thresholds are not especially strong predictors of language or academic achievement. Some investigators have found a relationship between auditory thresholds and language performance (e.g., Ching *et al.*, 2013; Sininger *et al.*, 2010; Stiles *et al.*, 2012), but that relationship is typically reported only when a very wide range of hearing levels are considered. When the hearing level is more tightly constrained, audiometric thresholds are poorer at explaining outcomes. For example, Davis *et al.* (1986) evaluated language and academic performance for 40 children with pure-tone average thresholds between 15 and 73 dB hearing level and found no evidence of a relationship between that performance and those thresholds. Other investigators have similarly failed to find an effect of degree of threshold shift on language outcomes when mild-to-moderate hearing loss only is considered (e.g., Bess *et al.*, 1998; Blair *et al.*, 1985; Moeller, 2000). Variability in the timing and quality of early intervention can certainly explain the tenuous relationship between auditory thresholds and language outcomes to some extent (Calderon and Naidu, 1998; Moeller, 2000; Nittrouer, 2010; Nittrouer and Burton, 2005; Yanbay *et al.*, 2014), but unexplained variability in language skills remains.

The current study emerged from the prospect that there may be explanatory factors related to the cochlear pathology itself that have not yet been fully explored. Even though the most obvious consequence of cochlear damage involves raised auditory thresholds, those thresholds cannot explain the variability observed in functional outcomes, even for adults (e.g., Dubno and Dirks, 1989; Souza and Tremblay, 2006; Walden and Walden, 2004). For example, Halpin and Rauch (2009) demonstrated that two patients with quite similar audiograms can perform very differently on tests of word recognition. A major auditory deficit, other than decreased sensitivity, that must be suspected as constraining language outcomes for both adults and children with hearing loss is diminished frequency selectivity, or resolution. Individuals with sensorineural hearing loss undoubtedly experience broadening of the auditory filters, but the extent of that broadening is hard to diagnose. Although the amount of broadening is related to the extent and location of the cochlear damage, variability nonetheless exists for losses in the mild to moderate range (e.g., Carney and Nelson, 1983; Florentine *et al.*, 1980; Hopkins and Moore, 2010; Souza *et al.*, 2012). Auditory filters are broader in individuals with cochlear pathology, largely because the filter-sharpening effects of the outer hair cells are reduced (e.g., Narayan *et al.*, 1998; Oxenham and Bacon, 2003). These broadened auditory filters have little effect on auditory functioning in quiet. Instead, they primarily take their toll on the abilities of those afflicted to function in noise, because these filters allow more noise to pass than do narrower filters.

[a]Author to whom correspondence should be addressed. Electronic mail: nittrouer.1@osu.edu

[b]Current address: Belmont University, 1900 Belmont Boulevard, Nashville, TN 37212.

 0001-4966/2015/137(4)/2004/11/$30.00

A few studies with adults have tried to measure this relationship by examining speech recognition under conditions of broadened auditory filters. Often, listeners with normal hearing participate in these experiments so the effects of broadened filters can be disassociated from the effects of loss of audibility, something that is hard to do when hearing-impaired listeners are the participants (e.g., Dubno and Dirks, 1989). In these studies, stimuli have commonly been static spectra, such as what might be found at syllable centers where vowel targets are located, an approach that seems reasonable because it allows examination of the robustness of the representation of these speech-related spectra in the face of noise. Typically it is observed, however, that the effects of broadened filters are minimal (Leek *et al.*, 1987; Turner and Van Tasell, 1984). It turns out that listeners with normal hearing are quite good at recovering static spectral shape with only small amplitude troughs between resonant peaks (i.e., formants). Similarly, small effects are observed when recognition is examined for consonants in consistent /ɑ/-consonant-/ɑ/ frames (Léger *et al.*, 2012).

But speech recognition does not consist of recovering sequences of static spectra. Instead, the spectrum that results from producing speech is constantly changing. Although more complex acoustically than discrete spectra, that pattern of change in spectral structure across time is highly informative when it comes to making judgments about linguistic units (e.g., Kewley-Port *et al.*, 1983). In fact, where vowel recognition is concerned, it has been shown that adults are better able to label syllable-medial vowels using the time-varying formant structure on either side of the target, rather than the target itself (e.g., Jenkins *et al.*, 1983; Strange *et al.*, 1983). Both because the time-varying patterns are acoustically complex and linguistically informative, it could be predicted that degraded spectral representations, as associated with broadened auditory filters, would be especially disruptive to the perceptual utility of this time-varying spectral structure. Support for that prediction has been found: When sentence-length materials are used, the effects of simulating broadened auditory filters range from minimal to large, as a function of extent of simulated smearing and signal-to-noise ratio (SNR) (e.g., ter Keurs *et al.*, 1992, 1993). To illustrate these effects, Fig. 1 from Baer and Moore (1993) displays scores for the recognition of key words in sentences by adults with normal hearing, using three levels of smearing at three SNRs. It appears in this figure that speech recognition diminished as the extent of smearing increased and as SNR decreased, with an interaction between these factors. This finding suggests that abnormally broad auditory filters have the greatest impact on continuous speech recognition, where listeners need to track the time-varying patterns of vocal-tract resonances.

It might be predicted that children would be especially affected by such broadened auditory filters, given that they have been found to depend highly upon just that kind of time-varying structure (i.e., formant movement). Compared to adults, children have demonstrated greater perceptual weighting of changing patterns of formant frequencies in decisions regarding phoneme identity. This finding has been observed for formant transitions that move from syllable-initial

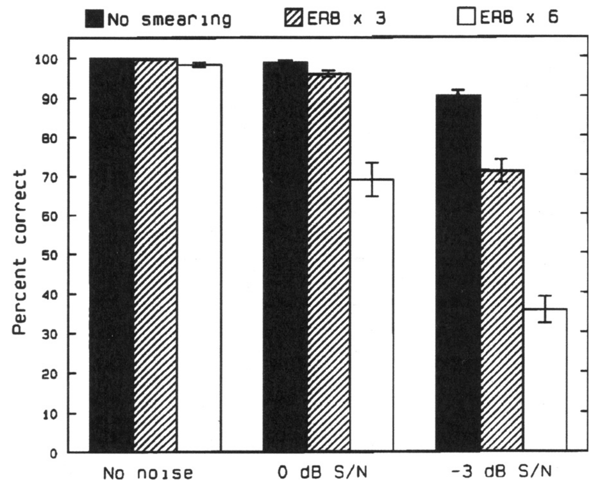

FIG. 1. Percent correct recognition of key words in sentences from adult listeners with normal hearing at three levels of smearing and three SNRs. ERB = equivalent rectangular bandwidth; S/N = signal-to-noise (ratio) (from Baer and Moore, 1993, reprinted with permission).

consonants to following vowels (e.g., Mayo *et al.*, 2003; Morrongiello *et al.*, 1984; Nittrouer and Studdert-Kennedy, 1987), as well as formant transitions that move from vowel nuclei to syllable-final consonants (e.g., Greenlee, 1980; Nittrouer, 2004; Wardrip-Fruin and Peach, 1984). In particular, it has been found that children are even more dependent on formant transitions than adults for the recognition of syllable-medial vowels (Nittrouer, 2007; Nittrouer and Lowenstein, 2014). At the level of the sentence, evidence of this strong selective attention to time-varying formant structure is provided by the fact that children are good at recognizing words in sentences processed to preserve only the first three formants as time-varying sine waves, and poorer at processing other spectrally degraded signals. Whereas adults have been found to recognize the same number of words correctly in sine-wave sentences as in four-channel noise-vocoded sentences, children recognize significantly more words correctly in sine-wave than in four-channel noise-vocoded sentences (Nittrouer *et al.*, 2009; Nittrouer and Lowenstein, 2010). Noise-vocoded signals can be viewed as an extreme form of spectral degradation, and other investigators have similarly observed that such degradation negatively impacts children's speech perception more than that of adults (e.g., Eisenberg *et al.*, 2000; Vongpaisal *et al.*, 2012). Findings such as these have led to speculation that attention to time-varying frequency structure is what children use to begin parsing the continuous, initially unanalyzable signals they hear into discrete lexical units (Nittrouer, 2006; Nittrouer *et al.*, 2009). According to this view, recurrent patterns of time-varying formant change in finite portions of the signal are recovered and constitute the early lexicon. Thus, the ability to recover this time-varying spectral structure is critical to the processes involved in language acquisition, such as early word learning. In turn, that idea means that the effects of broadened auditory filters associated with hearing loss could be predicted to disrupt early word learning especially strongly, and support for that prediction comes from studies showing difficulties specifically in word learning among children with mild-to-moderate hearing loss, compared to children with normal hearing (e.g., Gilbertson and Kamhi, 1995; Stelmachowicz *et al*., 2004). Again, however, outcomes in those studies for children with hearing loss did not vary with degree of threshold shift, raising the possibility that the extent of broadening of auditory filters may have an effect on language acquisition somewhat independently of auditory thresholds.

One experiment examined potential age-related differences in the effects of broadened filters using whispered speech (Nittrouer and Lowenstein, 2009). Although the direction and extent of formant movement is the same in whispered as in voiced speech, formants are broader. In that experiment, adults and children were asked to label the fricative in syllables consisting of synthetic noises from a /ʃ/-to-/s/ continuum followed by natural voiced or whispered vowels (excised from /ʃ/-vowel or /s/-vowel productions). Results showed that weighting coefficients for adults for the fricative noise spectrum and for the formant transition (appropriate for either /ʃ/ or /s/) did not vary depending on whether the vocalic portions were voiced or whispered. For 5- and 7-yr-olds, on the other hand, weighting coefficients for the fricative noise spectrum remained constant, but weighting of the formant transition was significantly less for whispered portions than for voiced portions. It was proposed that the effect arose because children had difficulty attending to those transitions when formant bandwidths were broad. In that experiment, a characteristic of speech production (i.e., whispering) led to those broader formants. In perception, characteristics of cochlear functioning could evoke the same effects.

A. Current study

The current study was motivated by two earlier, potentially related findings: First, it was motivated by the fact that children have been found to rely more than adults on time-varying formant structure in the speech signal. That finding has led to the suggestion that recurrent patterns of this time-varying structure are used by children to start parsing the largely uninterrupted speech signal that they hear into separate linguistic units, especially words. The second trend in earlier data that served to motivate this study was the well-replicated finding that neither degree of threshold shift nor factors related to treatment is able to completely explain the variability found in outcomes for children with mild-to-moderate hearing loss. It would be useful to identify the additional sources of variability so that appropriate treatments might be designed. In this study, it was specifically predicted that broadened auditory filters could diminish children's abilities to recover consistent patterns of time-varying spectral change. In turn, that decrement might explain some of the deficits in language outcomes observed for children with hearing loss, such as word-learning problems. The experiment described here was a first step to exploring that possibility.

Adults and children with normal hearing served as listeners in this experiment so that the effects of simulated broadened filters could be compared across listener age. Sentences served as stimuli because time-varying spectral structure was of particular interest. Earlier work showed that simulated broadened auditory filters have only small effects on recognition of static spectral structure for adults (Turner and Van Tasell, 1984). Precisely because children attend so strongly to time-varying speech structure, there was no reason to suspect that broadened auditory filters would have any stronger effects on recognition of static speech spectra for children. On the other hand, children might be especially hindered by broadened auditory filters in sentence recognition, where formant structure changes continuously across words.

Four hypotheses were tested in the current experiment. First, adults and children alike were expected to show diminished word recognition for sentences processed to simulate the broadened auditory filters associated with cochlear pathology, and then embedded in noise. Although the effect of listening with broadened filters has not been large or even necessarily significant when static spectra are used, it was hypothesized that listeners of all ages would show diminished word recognition for sentences when auditory filters were effectively broadened. This outcome was predicted by results of previous studies involving adults only (e.g., Baer and Moore, 1993; ter Keurs *et al*., 1992, 1993).

Second, it was predicted that the magnitude of this effect would be greater for children than for adults. This hypothesis arose specifically from results of earlier studies showing that children give strong perceptual attention to time-varying formant patterns. If this is the structure in the acoustic speech signal that is especially important to these young listeners, they should be especially hindered when it is degraded. Reason to suspect this outcome was provided by Nittrouer and Lowenstein (2009), where children showed diminished perceptual attention to whispered vowels, compared to voiced vowels.

To test a third prediction, the inverse of spectral smearing was performed: The sentence materials used in this experiment had their spectra enhanced, such that the amplitude differences between resonant peaks and valleys were increased over what they were in the unprocessed signals. Although this kind of processing has thus far largely been unsuccessful in improving recognition for speech in noise (Baer *et al*., 1993; Boers, 1980; Simpson *et al*., 1990), all of this work has been done with adults. It seemed important to try this manipulation with children, as well. Because children recognize sine-wave sentences disproportionately better than other kinds of degraded speech materials, it was considered possible that processing strategies deliberately narrowing the width of vocal-tract resonances might support improved speech recognition in noise for these listeners. A finding of that nature could have important implications for the design of auditory prostheses. To test these three predictions, young adults and children of two ages participated. By including two groups of children, potential developmental trends could be examined.

Two groups of the youngest and oldest listeners were included in this study, so that data could be collected at two SNRs for each of those age groups. The purpose of including these additional groups was two-fold: First, these groups allowed the exploration of the predictions described above, when listener groups were matched on performance level, rather than on SNR. Testing listeners in all three age groups at one SNR was expected to lead to large overall differences in performance; varying SNR across age groups in principled fashion was expected to make overall performance more equivalent. A second purpose to including these additional groups was that it permitted a test of the fourth hypothesis, which predicted an interaction between noise level and spectral smearing. Although Baer and Moore (1993) reported an interaction effect, that study was done with adults only, and outcomes were marred a bit by ceiling effects in a number of conditions, as shown in Fig. 1. Sentences were constructed in the current experiment to avoid ceiling effects, for any group.

II. METHOD

A. Participants

Forty adults (between the ages of 18 and 38), twenty 7-yr-olds (between the ages of 7 yrs; 0 months and 7 yrs; 3 months) and forty 5-yr-olds (between the ages of 4 yrs; 11 months and 5 yrs; 9 months) participated. All listeners were native speakers of American English, and none of the listeners (or their parents, in the case of children) reported any history of hearing or speech disorder. All listeners passed hearing screenings consisting of the pure tones of 0.5, 1, 2, 4, and 6 kHz presented at 20 dB hearing level to each ear separately. Parents reported that their children were free from significant histories of otitis media, defined as six or more episodes during the first three years of life. Children were given the Goldman Fristoe 2 Test of Articulation (Goldman and Fristoe, 2000) and were required to score at or better than the 30th percentile for their age in order to participate. The 7-yr-olds had a mean ranking of the 52nd percentile [standard deviation (SD) = 13] and the 5-yr-olds had a mean ranking of the 65th percentile (SD = 21). These scores indicate that the children in this study had normal articulation for their age. Adults were given the reading subtest of the Wide Range Achievement Test 4 (WRAT; Wilkinson and Robertson, 2006) and all demonstrated better than a 12th grade reading level. All listeners were also given the Expressive One-Word Picture Vocabulary Test (4th ed.) (EOWPVT; Martin and Brownell, 2011) and were required to achieve a standard score of at least 92 (30th percentile). The mean EOWPVT score for adults was 104 (SD = 9), which corresponds to the 61st percentile. The mean EOWPVT standard score for 5-yr-olds was 114 (SD = 10), corresponding to the 82nd percentile. The mean EOWPVT standard score for 7-yr-olds was 112 (SD = 10), which corresponds to the 79th percentile. These scores indicate that the adults had expressive vocabularies slightly above the mean of the normative sample used by the authors of the EOWPVT, and children had expressive vocabularies closer to 1 SD above the normative mean. Overall these screening instruments confirm that all participants had normal speech, language, and hearing abilities.

B. Equipment

All materials were recorded in a sound booth, directly onto the computer hard drive, via an AKG (Vienna, Austria) C535 EB microphone, a Shure (Niles, IL) M268 amplifier, and a Creative Laboratories (Singapore) Soundblaster soundcard. Perceptual testing took place in a sound booth, with the computer that controlled the experiment in an adjacent room. Stimuli were stored on a computer and presented through a Samson (Syosett, NY) headphone amplifier and AKG-K141 headphones. The hearing screening was done with a Welch Allyn (Skaneateles Falls, NY) TM262 audiometer and TDH-39 (Telephonics, Farmingdale, NY) headphones. All test sessions were video-recorded using a Sony (Japan) HDR-XR550V video recorder so that scoring could be done later. Participants wore Sony FM microphones that transmitted speech signals directly into the line input of the camera. This ensured good sound quality for all recordings.

C. Stimuli

Seventy-nine four-word sentences were used in this experiment; 4 for training and 75 for testing. These sentences were syntactically correct, but semantically anomalous. These sorts of sentences have been found to promote natural intonation and formant movement across words, without providing such strong top-down linguistic constraints that the effects of signal structure cannot be measured. Sentences of this type have often been used in the past, for the reasons described here (e.g., Boothroyd and Nittrouer, 1988). In earlier studies, children as young as 4 years of age have displayed context effects similar in magnitude to those of adults, when sentences with these simple constructions are used (Nittrouer and Boothroyd, 1990). That is, when language structures are within children's knowledge base, they are able to use that knowledge to facilitate recognition to the same extent as adults.

Fifty of the sentences used in testing and the four training sentences were taken from Nittrouer *et al.* (2014). An additional 25 sentences were created based on the low-predictability sentences generated by Stelmachowicz *et al.* (2000). All 79 sentences are listed in the Appendix. The sentences were recorded at a 44.1-kHz sampling rate with 16-bit digitization by an adult male speaker of American English. The sentences were down-sampled to 20 kHz before they were processed further. All of the stimuli were processed in two ways. First, spectra of the voiced portions were smeared to be half as sharp as the spectral envelopes of the original stimuli, meaning that excursions of spectral peaks and valleys were adjusted to be only half as far from the mean spectral slope as in the original signals. The second kind of processing done was that spectra of the voiced portions were enhanced to be twice as sharp as the spectral envelopes of the original stimuli, meaning that spectral peaks and valleys were adjusted to be twice as far from the mean spectral slope as in the original signals.

Before the processing was performed, the voiced signal portions needed to be located: Spectral smearing and enhancement were performed by manipulating the amplitude of individual harmonics, so it could be done only on those voiced portions. Boundaries between the voiceless and voiced signal portions were estimated by counting zero crossings in sequential 30-ms time frames, with 10-ms overlap. The expected number of zero crossings could be estimated based on talker gender. Voiced signal portions could then be identified as areas with regular zero crossings close to that estimate; unvoiced signal portions did not have regular crossings. Although a software algorithm originally estimated boundaries, all were checked by eye subsequently. Once voiced signal portions were identified, individual pitch periods were located, using the method described by Nittrouer *et al.* (2013). The fundamental frequency associated with each pitch period was derived by taking the inverse of the period. Next the amplitude of each pitch period across the sentence was measured and recorded. The individual harmonics within each pitch period were then put into separate bins. The mean amplitude function across bins was computed by fitting a logarithmic least-squares fit line, as shown in Fig. 2. Next, differences were computed between the amplitude in each bin and the value of that mean spectrum at the point where that particular bin was located. For the smeared signals, this difference was decreased by half in each bin, such that the amplitude in bins greater than the value of the mean spectrum at that location was decreased and the amplitude in bins less than the value of the mean spectrum at that location was increased. These bins, pre- and post-processing, are also shown in Fig. 2. For the enhanced signals, the amplitude of each bin relative to the mean spectrum was doubled. Once the harmonics in a pitch period were modified, the amplitude of that pitch period was adjusted to match its preprocessing value. Root-mean-square amplitude across all sentences was equalized.

The resulting stimuli were then embedded in speech-shaped noise that was based on the average long-term spectra across all of the stimuli. A different stretch of noise was used for each sentence, and each sentence was embedded in noise at each of three SNRs: −3, 0, and +3 dB. This resulted in three comparable sets of stimuli, at three different SNRs.

D. Procedures

All procedures were approved by the Institutional Review Board of the Ohio State University. After participants (or their parents, in the case of children) signed the consent form, the hearing screening was administered. The sentence materials were presented next. Half of the 5-yr-olds were presented with sentences at 0 dB SNR and half were presented with sentences at +3 dB SNR. Half of the adults were presented with sentences at 0 dB SNR and half were presented with sentences at −3 dB SNR. All 7-yr-olds heard the sentences at 0 dB SNR. Although it was not possible to predict what step size in SNR would result in precisely equivalent performance across age groups, these 3 dB steps were based on outcomes of earlier work (Nittrouer and Boothroyd, 1990), and were expected to achieve close-to-equivalent results.

For the sentence recognition task, the listener was seated across the table from the experimenter. The video camera was positioned to face the listener, who wore the FM transmitter. All responses were video and audio recorded. Practice sentences were presented before testing took place. For each practice sentence, the unprocessed version was played first, with no noise, and the listener was asked to repeat it. Then the unprocessed version embedded in noise was played, and the listener was asked to repeat it.

During testing, stimuli were presented in a single block of 75 sentences. Smeared, enhanced, and unprocessed stimuli were mixed, with the rule that for every group of three sentences, one of each processing type would be played in random order. This meant that no more than two items of one processing type could be presented in a row. The 25 sentences presented in each processing condition were determined by random selection for each listener before testing started. Each sentence was played once, and the listener repeated what was heard. Children moved a game piece along a five-space game board after every 15 sentences. This procedure served as a visible indicator of progress.

After the sentence recognition task was completed, the two screening tasks were administered: WRAT and EOWPVT for adults, and the Goldman-Fristoe and EOWPVT for 7- and 5-yr-olds. Although screening tools, these tasks were administered last so that all listeners would be optimally attentive during the sentence recognition task.

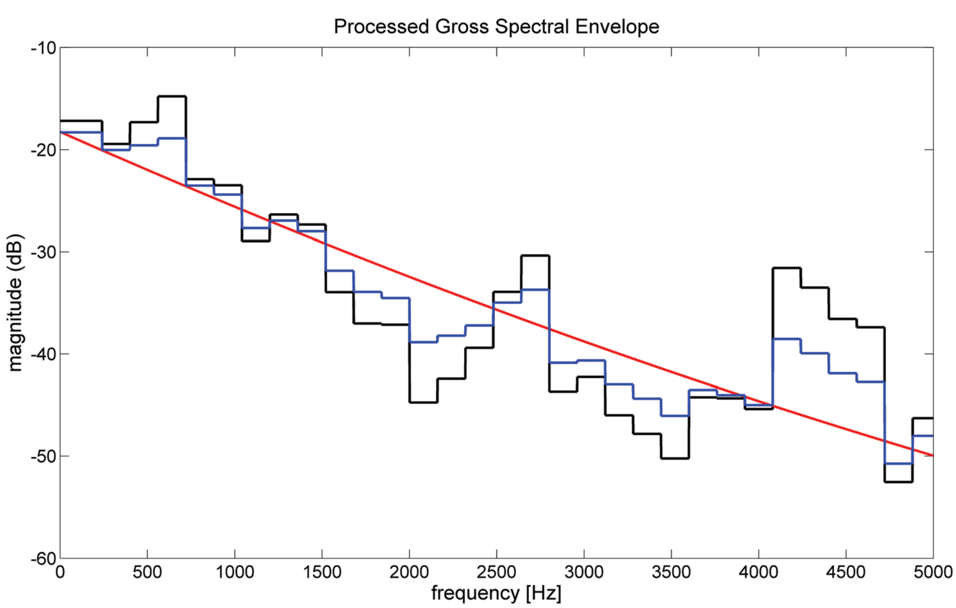

FIG. 2. Separate bins comprised of individual harmonics from one pitch period, preprocessing (black line) and post processing (blue line). The red line shows the mean spectral slope.

E. Scoring and analyses

The dependent measure was the number of words recognized correctly in each processing condition. All video-recorded responses were scored by the third author. In addition, the last author scored 25% of listeners' responses in each group (i.e., five in each group). Word-by-word agreement was computed between scores of the third and the last author for each listener (with two scores) as an index of inter-rater reliability. This was done by dividing the total number of agreements by 300, which was the total number of words in the 75 four-word sentences.

Although the dependent measure of interest was the number of words recognized correctly, the number of whole sentences recognized correctly was also recorded. The purpose of this procedure was to enable the computation of a metric of top-down linguistic context effects. Equation (1) from Boothroyd and Nittrouer (1988) was used for this purpose:

$$j = \log(p_s)/\log(p_w), \quad (1)$$

where p_s is the probability of recognizing whole sentences correctly and p_w is the probability of recognizing individual words correctly. Here, j represents the number of independent channels of information required for recognition, and is between 1 and the total number of words in the sentences. In this case, the smaller j is found to be, the greater the effect of syntactic constraints on recognition.

SPSS, version 21 was used for all analyses. Data were screened for normal distributions and homogeneity of variance across groups prior to conducting any statistical tests.

III. RESULTS

Across all samples that had been scored by two staff members, average agreement was 0.993 (SD = 0.005). This was considered good reliability, and scores from the third author were used in all further analyses. Data were found to satisfy the conditions of both normal distributions and homogeneity of variances across groups. A j factor was computed for each listener, and a one-way analysis of variance (ANOVA) was performed on these scores. No significant age effect was observed, so it was concluded that all listeners applied syntactic context effects to a similar extent. The mean j factor across all listeners was 3.24 (SD = 0.62). This is similar to what has been found for adults and children in earlier experiments (Boothroyd and Nittrouer, 1988; Nittrouer and Boothroyd, 1990).

In all analyses, a significance level of 0.05 was used. Nonetheless, in reporting outcomes, precise significance levels are given when $p < 0.10$; for $p > 0.10$, outcomes are reported simply as not significant.

Table I shows mean percent words correct for the three processing conditions, for each listener group separately. Mean scores from adults at both SNRs and 5-yr-olds at both SNRs are shown. Arcsine transformations were not applied to these data, because scores were not close to either 0% or 100% for any group of listeners. Consequently, statistical outcomes would not be affected by the use of these transformations.

TABLE I. Mean percent correct words for the three processing conditions, for each group at each SNR tested. Standard deviations are in parentheses, and overall means for each condition are at the bottom.

	Unprocessed	Enhanced	Smeared
5-year-olds (+3 dB)	65.5 (11.8)	56.8 (12.5)	55.8 (13.3)
5-year-olds (0 dB)	39.7 (11.3)	35.3 (9.8)	28.4 (11.4)
7-year-olds (0 dB)	59.6 (9.1)	53.3 (7.4)	46.2 (7.8)
Adults (0 dB)	74.5 (4.5)	71.9 (6.8)	63.2 (7.2)
Adults (−3 dB)	55.5 (6.5)	51.9 (7.8)	37.0 (8.0)
Mean	*59.0 (14.6)*	*53.9 (14.8)*	*46.1 (15.9)*

A. Equivalent SNR

Outcomes were first evaluated only for the listeners who were tested at 0-dB SNR. Scores for these listeners are in the three middle rows of Table I. Looking at scores for just the unprocessed sentences, it appears that the effects of noise decreased with increasing age. A one-way ANOVA performed on these scores revealed a significant age effect, $F(2,57) = 78.14$, $p < 0.001$, that was large in size, $\eta^2 = 0.73$. Furthermore, *post hoc* comparisons revealed significant differences among all groups, with $p < 0.001$ when Bonferroni adjustments for multiple comparisons were used. This large range of scores for what is effectively the baseline condition introduces a potential confound to any interpretation of the effects of spectral smearing or enhancement across age groups. Questions could be raised as to whether similar absolute differences in recognition scores between conditions represent equivalent effects when performance is so different overall. This concern is considered in Sec. III B.

For the three groups tested at 0-dB SNR, it appears that word recognition was poorer for both processed conditions—smeared and enhanced—than for the unprocessed condition, for all age groups. That finding is contrary to what had been predicted, which was that performance would be hindered for smeared stimuli, but improved for enhanced stimuli. However, the magnitude of the decrement in performance appears greater for the smeared than for the enhanced stimuli. To investigate these apparent effects, a two-way, repeated-measures ANOVA was performed, with processing condition as the repeated measure and age group as the within-subjects measure. Both main effects were significant: Processing condition, $F(2,114) = 69.67$, $p < 0.001$, $\eta^2 = 0.55$, and age, $F(2,57) = 116.44$, $p < 0.001$, $\eta^2 = 0.80$. However, the age × processing condition was not significant.

Tests of marginal effects (i.e., processing condition and age group) using Bonferroni adjustments for multiple comparisons showed that performance for each processing condition differed from each of the other processing conditions, and performance of each age group differed from that of each of the other age groups ($p < 0.001$ in all cases). Thus, performance was best for the unprocessed condition, followed by the enhanced condition, and finally by the smeared condition. That means that even though performance was diminished for the enhanced compared to the unprocessed stimuli, the effect was not as large as what was found for the smeared stimuli. Consequently, evidence was found to support the first hypothesis tested by this study: Listeners of all ages showed diminished recognition when auditory filters were effectively broadened. The third hypothesis described in Sec. I—that performance would be better when the spectrum was enhanced—was not supported by these data.

Another hypothesis tested in this study was that children would show a disproportionately larger effect of broadened auditory filters on speech recognition than adults. The lack of an age $\times$ processing condition interaction might be taken as evidence contrary to that hypothesis. However, Table I shows that children performed poorer than adults overall, and the finding of a significant main effect of age supported that observation. Consequently, a consistent difference in absolute scores between processing conditions for adults and children might represent a larger proportional decline in performance with processing for children. This concern was handled by trying to equalize overall performance across groups, which was done by running the youngest and oldest listeners at slightly different SNRs.

B. Equivalent performance

Next, recognition scores were evaluated when an attempt was made to equalize performance by varying SNR across age groups. In these analyses, outcomes were examined for a group of 5-yr-olds tested at +3 dB SNR, 7-yr-olds tested at 0 dB SNR, and adults tested at −3 dB SNR. Scores for these listeners are in the first, middle, and last rows of Table I, respectively. As an index of how similar performance was for these listener groups, scores for the unprocessed materials were compared. A one-way ANOVA performed on these scores revealed a significant age effect, $F(2,57) = 5.67$, $p = 0.006$, indicating that the attempt to equalize performance was not successful. However, the size of that effect was smaller than in the earlier analysis: here $\eta^2 = 0.17$ instead of 0.73, as found for the groups examined above. Moreover, the only *post hoc* comparison that was significant was for 5-yr-olds versus adults, $p = 0.004$ with a Bonferroni adjustment. Consequently, listeners in all groups performed more similarly overall than when a single SNR was used, making comparisons of outcomes across conditions based on listener age more appropriate in some sense than in the previous analysis. In this case, when a two-way, repeated-measures ANOVA was performed, both main effects were again significant: Processing condition, $F(2,114) = 84.64$, $p < 0.001$, $\eta^2 = 0.60$, and age, $F(2,57) = 8.78$, $p < 0.001$, $\eta^2 = 0.24$. This time, however, the age $\times$ processing condition was significant, as well, $F(4,114) = 7.43$, $p < 0.001$, $\eta^2 = 0.21$. Thus, here an age-related difference in the magnitude of the processing effect was observed.

In order to examine that difference across groups in the magnitude of the processing effect, difference scores were derived for the unprocessed condition compared to each of the processed conditions, and group means are shown in Table II. Outcomes of t tests for each comparison are also shown. It can clearly be seen that the decrement in performance associated with the smeared condition increases with increasing age, and the decrement in performance associated with the enhanced condition decreases with increasing age. In fact, for adults there is no statistically significant decrement in performance for the enhanced condition. One-way ANOVAs were performed on these difference scores. A significant age effect was obtained only for the Unprocessed-Smeared scores, $F(2,57) = 4.82$, $p = 0.012$, $\eta^2 = 0.15$, and *post hoc* comparisons revealed a significant difference only for 5-yr-olds versus adults, $p = 0.009$ with Bonferroni adjustments. Consequently, it might be concluded that an age effect was found for the impact of broadened auditory filters on speech recognition in noise, but the direction of effect was opposite to what was predicted: Adults showed larger effects than children.

TABLE II. Mean differences in percent words recognized correctly between processing conditions for each age group. Results of t tests are also shown. Degrees of freedom = 19 in all cases.

	Unprocessed - Smeared			Unprocessed - Enhanced		
	difference	t	p	difference	t	p
5-year-olds (+3 dB)	9.7 (7.5)	5.76	<0.001	8.7 (6.6)	5.86	<0.001
7-year-olds (0 dB)	13.5 (10.4)	5.80	<0.001	6.3 (8.9)	3.16	0.005
Adults (−3 dB)	18.5 (9.0)	9.2	<0.001	3.4 (8.5)	1.81	0.087

The fact that this age-related difference in the magnitude of the effect of spectral smearing was found only for adults versus 5-yr-olds is relevant because the age effect in performance for the unprocessed condition was also restricted to 5-yr-olds versus adults. That is, 5 - and 7-yr-olds performed statistically the same with the unprocessed stimuli, and 7-yr-olds and adults performed statistically the same. Thus, only where there was found a group difference in performance for unprocessed stimuli was a difference in the effect of spectral smearing found. Because there was a 6 dB difference in the SNR at which these two groups heard the stimuli, this outcome could have emerged from an interaction of SNR and spectral smearing, as suggested by the work of Baer and Moore (1993).

C. SNR interactions

The final set of analyses that was performed addressed this question of whether there was a processing condition $\times$ SNR interaction, as Baer and Moore (1993) had found. To achieve this goal, a three-way, repeated-measures ANOVA was performed on data from the two groups of 5-yr-olds and the two groups of adults. In this case, processing condition was the repeated measure, and both age and SNR served as between-subjects factors. In this case, SNR was assigned a binary code (i.e., more or less favorable), because absolute SNR was not matched across age groups. Results revealed that all three main effects were significant. First, processing condition was significant, $F(2,152) = 108.90$, $p < 0.001$, $\eta^2 = 0.60$, reflecting the finding that listeners performed best in the unprocessed condition, followed by the enhanced, and last by the smeared condition. Second, the effect of age was

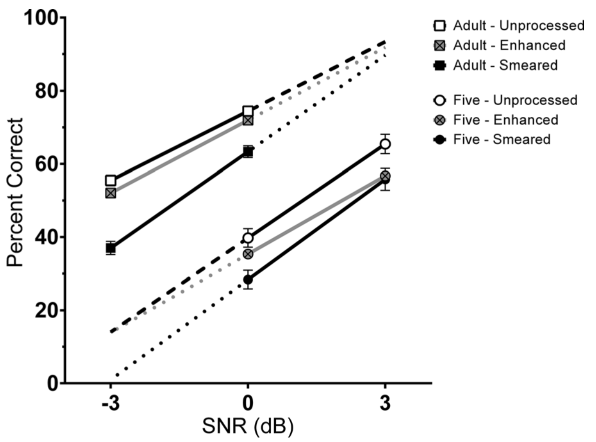

FIG. 3. Performance functions for adults and 5-yr-olds, with extrapolation.

significant, $F(1,76) = 39.99$, $p < 0.001$, $\eta^2 = 0.35$, reflecting the fact that adults performed better than children overall. Third, SNR was significant, $F(1,76) = 147.24$, $p < 0.001$, $\eta^2 = 0.66$, which indicated that listeners performed better at more favorable SNRs. In addition, two of the two-way interactions were significant: Processing condition $\times$ age, $F(2,152) = 10.37$, $p < 0.001$, $\eta^2 = 0.12$; and processing condition $\times$ SNR, $F(2,152) = 6.92$, $p = 0.001$, $\eta^2 = 0.08$. The first interaction replicates outcomes described in Sec. III B, that adults appear to have shown a larger effect of spectral smearing. The second interaction is of most interest here because it directly addresses the fourth hypothesis, that the effect of spectral smearing would be greater at poorer SNRs. Specifically, this significant two-way interaction term provides support for that prediction.

The three-way interaction of processing condition $\times$ SNR $\times$ age was not significant in this analysis, so it might be concluded that the interaction of processing condition $\times$ SNR was similar for adults and children. However, because adults and children were presented with stimuli at different SNRs, it seemed worthwhile to examine patterns of responding across SNRs for each group separately. To achieve that goal, mean performance for each group in each condition at each SNR was plotted and extrapolated so that functions were equally as extensive for the two groups. These functions are shown in Fig. 3, and slopes of these functions are shown in Table III. The patterns seen in Fig. 3 and Table III suggest that both adults and children showed a steeper drop in performance as SNR decreased for the smeared condition, compared to the unprocessed condition. This trend matches predictions from Baer and Moore (1993). However, the discrepancy in slope across these two conditions is greater for adults than for children. In fact, slopes of 5-yr-olds' functions for both the unprocessed and smeared conditions are more similar to adults' slope for the smeared condition than for their slope for the unprocessed condition. This finding suggests that children are simply more affected by noisy conditions than adults, regardless of auditory filter width. These outcomes also indicate that performance for unprocessed and spectrally smeared stimuli will attain equivalence sooner as SNR improves for adults than for children. Thus, adults with sensorineural hearing loss will be able to tolerate more noise than children with sensorineural hearing loss.

TABLE III. Slopes for performance functions shown in Fig. 3.

	Unprocessed	Enhanced	Smeared
Adults	6.33	6.67	8.73
5-yr-olds	8.60	7.17	9.13

Where slopes for the enhanced condition are concerned, adults had almost an identical slope in this condition to that found for the unprocessed condition. However, slope for the enhanced condition was the flattest of the three for 5-yr-olds, suggesting that—if done properly—spectral enhancement of the speech signal could facilitate speech recognition in noise for children. Although recognition scores were no better for the enhanced than for the unprocessed condition in this experiment, the finding of a flatter slope for the enhanced condition suggests that benefits would increase as SNR decreased, if signal enhancement were done in a way that improved performance generally in this condition.

IV. DISCUSSION

The study reported here was undertaken primarily to see if the broadened auditory filters associated with cochlear damage might affect speech recognition for children more than adults. The motivation underlying this work concerned the fact that language outcomes in children with hearing loss are variable, even when the amount of threshold shift and intervention factors are taken into account. Thus the possibility was considered that variability across children in the extent to which auditory filters might be broadened could account for some additional variability in outcomes. It has been suggested that time-varying patterns of formant frequencies play an especially important role in language acquisition, allowing children to extricate recurring sequences of these time-varying patterns in order to construct an early lexicon. If access to those time-varying spectral patterns is diminished due to broadened filters, a constraint on language acquisition could be predicted.

In total, four hypotheses tested specific predictions in this study. First, adults and children alike were predicted to show diminished recognition for sentences processed to simulate broadened auditory filters, and then embedded in noise. Second, it was predicted that the magnitude of this effect would be greater for children than for adults. The third prediction was that processing the sentence materials to enhance, rather than smear, the speech spectrum would benefit word recognition in noise. Finally, the fourth prediction tested by this study was that the effect of spectral smearing would be greater at poorer SNRs.

The first hypothesis was well supported by the data collected in this experiment: Adults and children alike showed diminished word recognition when the speech spectrum was smeared to simulate the broadened auditory filters imposed by cochlear damage. The second hypothesis, however, was not supported. In fact, at the SNRs used in this study, it was the adults who showed the largest effects. Of course, this outcome was observed when performance levels across age groups were matched as well as possible by presenting sentences at different SNRs across groups. As a result of that manipulation, a confound was introduced: Adults were listening at poorer SNRs than children.

No evidence was found to support the third hypothesis, that enhancing the speech spectrum would facilitate better recognition in noise. In fact, a trend seemingly opposite to that prediction was observed: Overall recognition scores were slightly diminished when the speech spectrum was enhanced, but this effect was significant only for children. However, failure to find evidence to support the hypothesis posed here may be due to the spectral enhancement method used. Manipulating the relative amplitude of separate harmonics surely had the desired effect where smearing was concerned: The spectrum as a whole was unquestionably flatter. When it came to enhancing the signal, however, this processing method may have had unintended consequences. Resonances other than those that arise in the oral vocal tract can be found in the speech spectrum. Some of these additional resonances are generated in the subglottal cavities, when the glottis is not sufficiently closed. Others may arise in the nasal cavity, when the velum is not sufficiently raised. These resonances, which are described by Stevens (1998), impart unique characteristics to the speech signal that can be associated with individual talkers. However, they generally have little, if any, linguistic significance. With the enhancement strategy implemented in the current experiment, these additional resonances would have been enhanced along with the linguistically significant resonances created in the oral cavity. The enhancement of these other resonances, which do not provide meaningful information, could have actually interfered with speech recognition. Thus, future efforts to provide appropriate spectral enhancement might involve extracting vocal-tract resonances first, and only enhancing those signal components. Alternatively, all spectral resonances that do not appear to be associated with the vocal tract could be attenuated prior to the enhancement process. Nonetheless, the shallower slope observed for 5-yr-olds' recognition function across SNRs for the enhanced compared to the unprocessed or smeared conditions suggests that signal enhancement might have beneficial effects at poor SNRs for children—if an appropriate method can be devised.

Regarding the fourth hypothesis, clear supporting evidence was obtained, but only for adults. Although the slopes of the performance functions across SNRs were greater for the smeared than for the unprocessed condition for both adults and 5-yr-olds (Table III), only adults showed larger differences in recognition between the unprocessed and smeared conditions at the poorer SNR (Table I). This finding helps to explain why speech recognition diminishes more precipitously with increasing noise levels for listeners with hearing loss, who experience some degree of broadened auditory filters, than for those with normal hearing.

In summary, the current study was undertaken to examine whether a kind of cochlear damage not previously well-studied in children might explain some of the variability in language outcomes observed for children with hearing loss. Results of the current study suggested that the specific problem under consideration—broadening of auditory filters arising from outer hair cell damage—might contribute in a meaningful way to this variability. Although greater effects of this spectral smearing were not observed for children than for adults, the toll to language learning might be especially great. This suggestion hinges on the fact that time-varying patterns of formant frequency movement are critical to early learning. Consequently, children might benefit more than adults from appropriate spectral enhancement, if such strategies can be developed.

ACKNOWLEDGMENTS

The authors wish to thank Jamie Kuess for writing the software to present stimuli. This work was supported by Grant No. R01 DC000633 from the National Institutes of Health, National Institute on Deafness and Other Communication Disorders.

APPENDIX: SENTENCES USED IN TESTING

1. Practice sentences

P1 Ducks teach sore camps.
P2 Find girls these clouds.
P3 Cooks run in brooms.
P4 Great shelf needs tape.

2. Test sentences

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

30 Cars jump from fish.
31 They turn small trees.
32 Trucks drop sweet dust.
33 Let their flood hear.
34 Long kids stay back.
35 Guys tell loud meat.
36 Thin books look soft.
37 Snow smells more tough.
38 Cups kill fat leaves.
39 Blue chairs speak well.
40 Slow dice buy long.
41 Lead this coat home.
42 Pink chalk bakes phones.
43 Shy laws have keys.
44 High bears move holes.
45 Call her wing guide.
46 Four rats kick warm.
47 Soft rocks taste red.
48 Cold worms have toys.
49 Fan spells large toy.
50 Jobs get thick hay.
51 Clocks hold rough cows.
52 Brown nights dance more.
53 Now straws need cheese.
54 Please shine some clowns.
55 Most birds knock tea.
56 Large food hikes loose.
57 Black frogs bring Mom.
58 Rich men might pop.
59 Take splash from her.
60 Floors hug dull juice.
61 Bread drinks hot farms.
62 Jokes fall on tails.
63 Nice bugs itch far.
64 He was the rain.
65 Dad bites dry bowls.
66 Her hill could bike.
67 Rude pigs drank shirts.
68 Strange nails taste dark.
69 All boys are paint.
70 We fly like chairs.
71 Dull socks wag off.
72 It poked sore trains.
73 Great gum hurts jam.
74 Tall pools scare lamps.
75 Good boats tease pants.

Baer, T., Moore, B. C., and Gatehouse, S. (**1993**). "Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times," J. Rehabil. Res. Dev. **30**, 49–72.

Baer, T., and Moore, B. C. J. (**1993**). "Effects of spectral smearing on the intelligibility of sentences in noise," J. Acoust. Soc. Am. **94**, 1229–1241.

Bess, F. H., Dodd-Murphy, J., and Parker, R. A. (**1998**). "Children with minimal sensorineural hearing loss: Prevalence, educational performance, and functional status," Ear Hear. **19**, 339–354.

Blair, J. C., Peterson, M. E., and Viehweg, S. H. (**1985**). "The effects of mild sensorineural hearing loss on academic performance of young school age children," Volta Rev. **87**, 87–93.

Boers, P. M. (**1980**). "Formant enhancement of speech for listeners with sensorineural hearing loss," IPO Ann. Prog. Rep. **15**, 21–28.

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

Calderon, R., and Naidu, S. (**1998**). "Further support for the benefits of early identification and intervention for children with hearing loss," Volta Rev. **100**, 53–84.

Carney, A. E., and Nelson, D. A. (**1983**). "An analysis of psychophysical tuning curves in normal and pathological ears," J. Acoust. Soc. Am. **73**, 268–278.

Ching, T. Y., Dillon, H., Marnane, V., Hou, S., Day, J., Seeto, M., Crowe, K., Street, L., Thomson, J., Van Buynder, P., Zhang, V., Wong, A., Burns, L., Flynn, C., Cupples, L., Cowan, R. S., Leigh, G., Sjahalam-King, J., and Yeh, A. (**2013**). "Outcomes of early- and late-identified children at 3 years of age: Findings from a prospective population-based study," Ear Hear. **34**, 535–552.

Davis, J. M., Elfenbein, J., Schum, R., and Bentler, R. A. (**1986**). "Effects of mild and moderate hearing impairments on language, educational, and psychosocial behavior of children," J. Speech Hear. Dis. **51**, 53–62.

Dubno, J. R., and Dirks, D. D. (**1989**). "Auditory filter characteristics and consonant recognition for hearing-impaired listeners," J. Acoust. Soc. Am. **85**, 1666–1675.

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.

Florentine, M., Buus, S., Scharf, B., and Zwicker, E. (**1980**). "Frequency selectivity in normally-hearing and hearing-impaired observers," J. Speech Hear. Res. **23**, 646–669.

Gilbertson, M., and Kamhi, A. G. (**1995**). "Novel word learning in children with hearing impairment," J. Speech Hear. Res. **38**, 630–642.

Goldman, R., and Fristoe, M. (**2000**). *Goldman Fristoe 2: Test of Articulation* (American Guidance Service, Inc., Circle Pines, MN), pp. 1–146.

Greenlee, M. (**1980**). "Learning the phonetic cues to the voiced-voiceless distinction: A comparison of child and adult speech perception," J. Child Lang. **7**, 459–468.

Halpin, C., and Rauch, S. D. (**2009**). "Clinical implications of a damaged cochlea: Pure tone thresholds vs information-carrying capacity," Otolaryngol. Head Neck Surg. **140**, 473–476.

Hopkins, K., and Moore, B. C. (**2010**). "The importance of temporal fine structure information in speech at different spectral regions for normal-hearing and hearing-impaired subjects," J. Acoust. Soc. Am. **127**, 1595–1608.

Jenkins, J. J., Strange, W., and Edman, T. R. (**1983**). "Identification of vowels in 'vowel-less' syllables," Percept. Psychophys. **34**, 441–450.

Kewley-Port, D., Pisoni, D. B., and Studdert-Kennedy, M. (**1983**). "Perception of static and dynamic acoustic cues to place of articulation in initial stop consonants," J. Acoust. Soc. Am. **73**, 1779–1793.

Leek, M. R., Dorman, M. F., and Summerfield, Q. (**1987**). "Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **81**, 148–154.

Léger, A. C., Moore, B. C., Gnansia, D., and Lorenzi, C. (**2012**). "Effects of spectral smearing on the identification of speech in noise filtered into low- and mid-frequency regions," J. Acoust. Soc. Am. **131**, 4114–4123.

Martin, N., and Brownell, R. (**2011**). *Expressive One-Word Picture Vocabulary Test (EOWPVT)*, 4th ed. (Academic Therapy Publications, Inc., Novato, CA), pp. 1–99.

Mayo, C., Scobbie, J. M., Hewlett, N., and Waters, D. (**2003**). "The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception," J. Speech Lang. Hear. Res. **46**, 1184–1196.

Moeller, M. P. (**2000**). "Early intervention and language development in children who are deaf and hard of hearing," Pediatrics **106**, E43.

Morrongiello, B. A., Robson, R. C., Best, C. T., and Clifton, R. K. (**1984**). "Trading relations in the perception of speech by 5-year-old children," J. Exp. Child. Psychol. **37**, 231–250.

Narayan, S. S., Temchin, A. N., Recio, A., and Ruggero, M. A. (**1998**). "Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae," Science **282**, 1882–1884.

Nittrouer, S. (**2004**). "The role of temporal and dynamic signal components in the perception of syllable-final stop voicing by children and adults," J. Acoust. Soc. Am. **115**, 1777–1790.

Nittrouer, S. (**2006**). "Children hear the forest," J. Acoust. Soc. Am. **120**, 1799–1802.

Nittrouer, S. (**2007**). "Dynamic spectral structure specifies vowels for children and adults," J. Acoust. Soc. Am. **122**, 2328–2339.

Nittrouer, S. (**2010**). *Early Development of Children with Hearing Loss* (Plural Publishing, San Diego, CA), pp. 1–345.

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., and Lowenstein, J. H. (**2009**). "Does harmonicity explain children's cue weighting of fricative-vowel syllables?," J. Acoust. Soc. Am. **125**, 1679–1692.

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., and Lowenstein, J. H. (**2014**). "Dynamic spectral structure specifies vowels for adults and children," Lang. Speech **57**, 487–512.

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.

Nittrouer, S., Lowenstein, J. H., and Tarr, E. (**2013**). "Amplitude rise time does not cue the /bɑ/-/wɑ/ contrast for adults or children," J. Speech Lang. Hear. Res. **56**, 427–440.

Nittrouer, S., and Studdert-Kennedy, M. (**1987**). "The role of coarticulatory effects in the perception of fricatives by children and adults," J. Speech Hear. Res. **30**, 319–329.

Nittrouer, S., Tarr, E., Bolster, V., Caldwell-Tarr, A., Moberly, A. C., and Lowenstein, J. H. (**2014**). "Low-frequency signals support perceptual organization of implant-simulated speech for adults and children," Int. J. Audiol. **53**, 270–284.

Oxenham, A. J., and Bacon, S. P. (**2003**). "Cochlear compression: Perceptual measures and implications for normal and impaired hearing," Ear Hear. **24**, 352–366.

Simpson, A. M., Moore, B. C., and Glasberg, B. R. (**1990**). "Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners," Acta Otolaryngol. Suppl. **469**, 101–107.

Sininger, Y. S., Grimes, A., and Christensen, E. (**2010**). "Auditory development in early amplified children: Factors influencing auditory-based communication outcomes in children with hearing loss," Ear Hear. **31**, 166–185.

Souza, P., Wright, R., and Bor, S. (**2012**). "Consequences of broad auditory filters for identification of multichannel-compressed vowels," J. Speech Lang. Hear. Res. **55**, 474–486.

Souza, P. E., and Tremblay, K. L. (**2006**). "New perspectives on assessing amplification effects," Trends Amplif. **10**, 119–143.

Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R. W., and Pittman, A. L. (**2000**). "The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children," J. Speech Lang. Hear. Res. **43**, 902–914.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., and Lewis, D. E. (**2004**). "Novel-word learning in children with normal hearing and hearing loss," Ear Hear. **25**, 47–56.

Stevens, K. N. (**1998**). *Acoustic Phonetics* (The MIT Press, Cambridge, MA), pp. 1–607.

Stiles, D. J., Bentler, R. A., and McGregor, K. K. (**2012**). "The Speech Intelligibility Index and the pure-tone average as predictors of lexical ability in children fit with hearing aids," J. Speech Lang. Hear. Res. **55**, 764–778.

Strange, W., Jenkins, J. J., and Johnson, T. L. (**1983**). "Dynamic specification of coarticulated vowels," J. Acoust. Soc. Am. **74**, 695–705.

ter Keurs, M., Festen, J. M., and Plomp, R. (**1992**). "Effect of spectral envelope smearing on speech reception. I," J. Acoust. Soc. Am. **91**, 2872–2880.

ter Keurs, M., Festen, J. M., and Plomp, R. (**1993**). "Effect of spectral envelope smearing on speech reception. II," J. Acoust. Soc. Am. **93**, 1547–1552.

Turner, C. W., and Van Tasell, D. J. (**1984**). "Sensorineural hearing loss and the discrimination of vowel-like stimuli," J. Acoust. Soc. Am. **75**, 562–565.

Vongpaisal, T., Trehub, S. E., Glenn Schellenberg, E., and van Lieshout, P. (**2012**). "Age-related changes in talker recognition with reduced spectral cues," J. Acoust. Soc. Am. **131**, 501–508.

Walden, T. C., and Walden, B. E. (**2004**). "Predicting success with hearing aids in everyday living," J. Am. Acad. Audiol. **15**, 342–352.

Wardrip-Fruin, C., and Peach, S. (**1984**). "Developmental aspects of the perception of acoustic cues in determining the voicing feature of final stop consonants," Lang. Speech **27**, 367–379.

Wilkinson, G. S., and Robertson, G. J. (**2006**). *The Wide Range Achievement Test (WRAT)*, 4th ed. (Psychological Assessment Resources, Lutz, FL), pp. 1–494.

Yanbay, E., Hickson, L., Scarinci, N., Constantinescu, G., and Dettman, S. J. (**2014**). "Language outcomes for children with cochlear implants enrolled in different communication programs," Cochlear Implants Int. **15**, 121–135.