Perceptual coherence in listeners having longstanding childhood hearing losses, listeners with adult-onset hearing losses, and listeners with normal hearing

Andrea Pittman[a),b)]
Arizona State University

a)*A portion of the data were collected while the author was affiliated with Boys Town National Research Hospital.*

b)*Electronic mail: andrea.pittman@asu.edu*

Abstract

Perceptual coherence, the process by which the individual elements of complex sounds are bound together, was examined in adult listeners with longstanding childhood hearing losses, listeners with adult-onset hearing losses, and listeners with normal hearing. It was hypothesized that perceptual coherence would vary in strength between the groups due to their substantial differences in hearing history. Bisyllabic words produced by three talkers as well as comodulated three-tone complexes served as stimuli. In the first task, the second formant of each word was isolated and presented for recognition. In the second task, an isolated formant was paired with an intact word and listeners indicated whether or not the isolated second formant was a component of the intact word. In the third task, the middle component of the three-tone complex was presented in the same manner. For the speech stimuli, results indicate normal perceptual coherence in the listeners with adult-onset hearing loss but significantly weaker coherence in the listeners with childhood hearing losses. No differences were observed across groups for the non-speech stimuli. These results suggest that perceptual coherence is relatively unaffected by hearing loss acquired during adulthood but appears to be impaired when hearing loss is present in early childhood.

Keywords

43.71.Ky; 43.71.Ft; 43.71.-k

I. INTRODUCTION

Speech is a stream of complex, dissimilar acoustic elements that are interpreted as orderly and meaningful events by experienced listeners. The process by which these disparate elements are bound together and perceived as speech has been given the term "perceptual coherence". Both phonemic and auditory cues within the speech signal have been shown to promote perceptual coherce. At the level of the phoneme, the listener appears to bind (or cohere) the various acoustic components within the phoneme during perception (Best et al., 1989; Remez & Rubin, 1990; Carrell & Opie, 1992; Barker & Cooke, 1999). Remez et al. (2001) demonstrated the coherence of acoustic components using simple sine-wave speech. In sine-wave speech, the formants that comprise a word or phrase are replicated with puretones that

Corresponding Author: Andrea Pittman, Ph.D. Arizona State University Dept of Speech and Hearing Science P.O. Box 870102 Tempe, Arizona 85287-0102.

vary in frequency and amplitude over time. Listeners in this study were first asked to detect differences between pairs of isolated second formants and then to indicate if the isolated formants occurred within synthetic words. Although the listeners were able to differentiate the isolated formants well (d-prime ~3), they were unable to detect the presence of a specific formant within a word (d-prime ~0). The results suggest that the process of binding the acoustic elements together resulted in the percept of one event rather than several individual events.

Perceptual coherence appears to operate in the same manner as comodulation masking release (Hall et al., 1984; McFadden, 1987; Grose & Hall, 1993). During comodulation masking release, the threshold of detection for a tone presented in an amplitude modulated band of noise is improved (i.e., threshold is reduced) by adding another amplitude modulated band of noise in a different frequency region, provided that the modulation rates are the same. It is thought that the listener perceives the two noise bands as one auditory event and thus perceives the tone separately, improving the threshold. During speech perception, the components of speech are thought to group together as one auditory event and stand apart from other simultaneous acoustic events (Bregman, 1990).

Gordon (1997) demonstrated this aspect of coherence using synthetic representations of the vowels /E/ and /I/. The first formants were centered at 625 and 375 Hz, respectively. The second and third formants were identical for both vowels and occurred at 2200 and 2900 Hz, respectively. All of the formants were amplitude modulated (or comodulated) at 125 Hz. Because the two high-frequency formants were the same for both vowels, the low-frequency formant provided the only unique information with which to identify either vowel. The stimuli were mixed with low-pass noise (corner frequency: 1000 Hz) that overlapped the first formant only. The listeners in this study were asked to indicate which of the two vowels was presented in each trial. The signal-to-noise ratio necessary to distinguish the two vowels was measured with and without the addition of the ambiguous higher-frequency formants. The results showed a small but significant reduction in signal-to-noise ratio (3 dB) when the high-frequency formants were added to the speech signal. These results suggest that the comodulated formants cohered to become one auditory event, improving the threshold of detection in noise.

Similarly, Carrell and Opie (1992) showed that intelligibility of sine-wave speech increased when the sinusoids were amplitude comodulated at rates similar to the fundamental frequency of adult talkers. In the third of three experiments, they comodulated the sinusoidal formants of four sentences at 50, 100, and 200 Hz and presented them for recognition to normally hearing listeners. They also presented the sentences with no comodulation. Because comodulation masking release is strongest at low modulation rates and weakest at high modulation rates, the authors hypothesized that intelligibility of the sentences would be best for the 50 Hz modulation rate and poorest for the 200 Hz modulation rate. The intelligibility results were consistent with their hypothesis and suggested that the principles that govern comodulation masking release are also involved in speech perception. That is, the talker's fundamental frequency contributes to perceptual coherence and that lower fundamental frequencies may promote stronger coherence.

Although studies in normally hearing adults have provided important evidence regarding the nature of perceptual coherence and the principles that govern it, additional insight may be gained from examination of the underdeveloped perceptual system (children) or the disordered system (listeners with hearing loss). In the only study of perceptual coherence in children to date, Nittrouer and Crowther (2001) used a categorical perception paradigm to examine the developmental course of phonemic coherence. Adults and children (5 and 7 years of age) were asked to indicate if two stimuli were the same or different. The second of the two stimuli contained gap duration and formant frequency cues that were consistent with naturally produced speech or conflicted with it. The consistent and conflicting cues were intended to

influence the listener's decision regarding the similarity between the two stimuli. The responses of the youngest children were least influenced by the consistent and conflicting cues contained in the stimuli suggesting stronger perceptual coherence in the 5-year-old children compared to the 7-year-olds and adults. Their results suggest that young children may learn to perceive speech by overcoming the effects of coherence rather than developing the ability to bind acoustic elements together. That is, children may learn to detect the subtle, but important, acoustic cues that lead to speech perception (e.g., the third formant distinction between /la/ and /ra/) rather than learning which cues to group together.

Unfortunately, there are no data regarding perceptual coherence in the impaired auditory system. However, several studies have examined the effect of hearing loss for a number of related auditory processes. For example, comodulation masking release has been shown to be significantly reduced in listeners with hearing loss relative to listeners with normal hearing (Hall & Grose, 1989; Hall & Grose, 1994; Eisenberg et al., 1995). Also, listeners with hearing loss have been shown to have difficulty organizing non-speech stimuli for the purpose of analytic listening (i.e., hearing out individual parts of a complex sound) (Grose & Hall, 1996; Rose & Moore, 1997; Kidd et al., 2002). Although the poor performance of listeners with hearing loss has been attributed to peripheral factors associated with hearing loss (e.g., abnormally wide auditory filters, poor temporal resolution, poor frequency selectivity), the contribution of these factors does not appear to be substantial and suggests that other factors may also contribute to coherence. Therefore, it may be informative to examine the perceptual coherence of listeners with adult-onset hearing losses relative to those having longstanding childhood hearing losses. The ability to organize the acoustic components of speech and non-speech stimuli may differ across listeners for whom perceptual coherence developed in the presence of normal hearing or in the presence of hearing loss.

The purpose of the current investigation was to examine the perceptual coherence of listeners with childhood hearing losses and listeners with adult-onset hearing losses relative to that of listeners with normal hearing. For this study, a version of the paradigm employed in Remez et al. (2001) was used (described above). This paradigm required the listener to detect individual acoustic components within speech and non-speech stimuli. Poor performance on this task would suggest relatively strong perceptual coherence whereas good performance would suggest relatively weak perceptual coherence. It was hypothesized that for both types of stimuli, the perceptual coherence of the listeners with adult-onset hearing losses would be stronger than normal due to their inability to perceive the subtle acoustic cues of speech. It was also hypothesized that the perceptual coherence of the listeners with childhood hearing loss would be greater than that of the listeners with adult-onset hearing loss due to deficits associated with life-long hearing loss as well as their inability to perceive the subtle acoustic cues of speech.

II. METHODS

A. Participants

Ten normal-hearing listeners, 10 listeners with adult-onset hearing losses (A-HL), and 10 listeners with longstanding, childhood hearing losses (C-HL) participated in this study. All of the listeners with hearing loss had bilateral, symmetrical, sensorineural hearing losses. The age and hearing levels for each of the listeners with hearing loss is given in Table I. The average hearing levels (and standard error) of the listeners with normal hearing are also given. On average, the thresholds of the listeners with A-HL were within the range of normal at 125 and 250 Hz and increased to ~70 dB SPL at frequencies >1000 Hz. The listeners with C-HL had equally poor thresholds across frequency (~60 dB SPL).

B. Stimuli

Three different stimuli were used in this study: isolated second formants, naturally produced words, and time-varying sinusoids. The stimuli were fashioned after those used in Remez et al. (2001) with the exception that the speech stimuli were naturally produced and all of the stimuli were longer in duration. The speech stimuli were selected from an original stimulus set containing digital recordings of thirty bisyllabic words produced by an adult male, an adult female, and a 7-year-old child. The words consisted of voiced consonants and vowels (sonorants). The words produced by each talker were concatenated and saved to separate files (one per talker). The stimuli were then subjected to a pitch extraction algorithm in Matlab (YIN: de Cheveigne & Kawahara, 2002). On average, the male, female, and child talkers had fundamental frequencies of 96, 180, and 227 Hz, respectively. These stimuli provided a range of fundamental frequencies which was expected to influence perceptual coherence in a systematic fashion. Specifically, coherence was expected to be best for the male talker (lowest fundamental frequency) and poorest for the child talker (highest fundamental frequency).

The second formant in each word was isolated using custom laboratory software written in Matlab. The program allowed the user to visually inspect the spectrogram and manually trace the second formant. The trace was then converted into a band-pass filter having a selectable bandwidth. For these stimuli, a filter with a bandwidth of 250 Hz (at the 3 dB down point) was applied throughout the length of the formant. The stimuli were presented for identification to a separate group of 10 normal-hearing adults. The words *ruler*, *lily*, *oily*, *hourly*, *rally*, *relay*, *layer*, *royal*, and *willow* could not be identified by any of the listeners based on the acoustic information provided by the second formant alone and were selected as stimuli for the study. Figure 1 shows the spectrograms for the second formants extracted from the nine stimulus words produced by the female talker. The isolated formants varied in frequency over time between 750 and 3500 Hz across talkers.

A series of simple time-varying sinusoids also were created. The sinusoids were amplitude modulated at 100 Hz using a triangular wave with a 50% duty cycle. The sinusoids were arranged into 3-tone complexes having a high, mid, and low-frequency component. The on- and off-set frequencies of each component in the complexes are listed in Table II. Note that the high-frequency component was the same for all stimuli and consisted of on- and off-set frequencies of 3500 and 4000 Hz, respectively. Likewise, the low-frequency component was the same for all stimuli and consisted of on- and off-set frequencies of 1000 and 500 Hz, respectively. The on- and off-set frequencies of the mid-frequency component varied from 1500 to 3000 Hz in 6 steps of 200 to 500 Hz.

C. Amplification Parameters

Prior to testing, hearing thresholds were obtained from each listener. The stimuli were frequency shaped for the listeners with hearing loss according to the Desired Sensation Level fitting algorithm (DSL i/o version 4.1) (Seewald, Cornelisse, Ramji, Sinclair, Moodie, & Jamieson, 1997). The amplification parameters provided levels sufficient to perceive average conversational speech (65 dB SPL). An estimate of stimulus audibility was calculated for each listener using the speech intelligibility index (SII). The importance function for nonsense syllables was used (ANSI, 1997). On average, the SII for the listeners with A-HL and C-HL was 0.84 (SD=0.05) and 0.73 (SD=0.23), respectively, which is sufficient for maximum perception of speech. Figure 2 shows the average long-term speech spectra (dashed lines) relative to the average hearing levels (solid lines) for the listeners with A-HL and C-HL in the upper and lower panels, respectively. Note that sufficient amplification was provided to ensure audibility of formants occurring at or below 4000 Hz for all listeners. The hatched area in each panel represents the range of thresholds for the listeners with normal hearing as defined by ±1

standard deviation around the mean for both the right and left ears. All measures were referenced to a 6-cm^3 coupler.

D. Procedures

Each listener participated in three tasks. In the first task, the listeners' ability to perceive a word from the second formant alone was determined. This task was necessary to confirm that the listeners perceived the second formants as acoustic stimuli rather than as meaningful speech. On each trial, an isolated second formant was presented and the listener verbally responded with the word he/she perceived, if any. The listeners were not provided with a list of words from which to choose their responses nor were they given any feedback. Each second formant was presented once for a total of 27 trials (9 formants × 3 talkers).

The second (and primary) task assessed the listener's ability to perceive a specific auditory object (the second formant) within a speech stimulus (the intact word). Each trial consisted of an isolated formant, an 800 ms silent gap, and an intact word. Figure 3 shows the waveform and spectrogram of one stimulus produced by the female talker. In this example, the second formant was extracted from the word *relay* and paired with the word *rally*. For each trial the listener reported whether or not the formant was contained within the word. The listeners indicated their response by selecting 'yes' or 'no' buttons on a touch-screen monitor. Feedback was provided for correct responses.

To minimize the number of stimuli while maximizing the difficulty of the task, the isolated formants were paired with words having similar second formant characteristics. To achieve this, the formants were ordered according to offset frequency and overall morphology. This order is given in Table III (as well as in Figure 1) and shows the manner in which the second formants (rows) were paired with the words (columns). Half of the trials contained a second formant and intact word which neighboured the original word ('no' trials). The remaining trials contained stimuli with the same second formants ('yes' trials). The words *ruler* and *willow* were not paired with neighbouring formants because they comprised the end points of the continuum. Admittedly, the progression from one formant to the next could have been accomplished through several different arrangements; however the perceptual differences between these formants was quite subtle. This task consisted of a total of 168 trials (3 talkers × 2 repetitions of the 14 'no' stimuli + 4 repetitions of the 7 'yes' stimuli).

The third task assessed the listener's ability to perceive a single amplitude modulated sinusoid within a three-tone complex of sinusoids. Because this task contained non-speech stimuli, coherence under these conditions may suggest that the fundamental acoustic characteristic of frequency modulation is a significant contributor to the coherence of speech. As with the speech stimuli, a mid-frequency sinusoid was paired with a three-tone complex containing the same mid-frequency sinusoid or that of a neighbouring complex. Figure 4 shows the waveform and spectrogram of one stimulus pair. Table IV shows the stimulus pairs in their respective 'yes' and 'no' categories. The procedures for this task were the same as those for the previous task. After hearing an isolated sinusoid followed by a three-tone complex, the listener indicated whether or not the isolated tone was contained in the complex. This task consisted of 80 trials (5 repetitions of the 8 'no' stimuli + 10 repetitions of the 4 'yes' stimuli).

All testing took place in an audiometric test suite. The stimuli were presented binaurally in quiet under earphones (Sennheiser, HD25). Testing required approximately 2 hours and each listener was paid for his/her time.

III. RESULTS

A. Task 1: Perception of words from isolated second formants

In general the listeners did not perceive the isolated second formants as speech. Recall that this task consisted of 27 trials (1 repetition × 9 formants × 3 talkers). On average, the listeners with normal hearing, A-HL, and C-HL correctly identified the word from which the isolated formant originated only 4, 1, and 3% of the trials, respectively. The highest score achieved was 11% (3 of 27 trials) by two normally hearing listeners and two listeners with C-HL. However, most of the other listeners (17 of 30) did not identify any of the isolated second formants as the words from which they originated. These results suggest that the listeners perceived the isolated formants as acoustic stimuli rather than speech. This result was important because it demonstrated that in the primary task the listeners were not simply comparing two speech stimuli.

B. Task 2: Coherence of real speech

Because the average performance of a listener may be biased toward one response in a two-alternative forced choice paradigm (yes/no), each listener's sensitivity to the 'yes' and 'no' trials was calculated and represented as a value of d-prime (Marshall & Jesteadt, 1986). Specifically, a listener's correct responses on 'yes' trials (hits) were subtracted from the portion of time he/she responded incorrectly to 'no' trials (false alarms). The d-prime value was then used to calculate the listener's maximized performance in the absence of a response bias. On average the maximized performance for each group was adjusted by less than 1 percentage point suggesting that these listeners demonstrated little or no response bias.

Figure 5 displays the average maximized performance (bars re: left axis) and d-prime (filled circles re: right axis) for the primary task as a function of group. The error bars represent ±1 standard error. In general, the performance and d-prime values for the listeners with C-HL were higher than that of the listeners with normal hearing or with A-HL. A repeated measures ANOVA of the d-prime data revealed a significant main effect of group ($F_{2,27} = 4.862$; p=0.016). Tukey's HSD post-hoc analyses revealed that, on average, the d-prime values of the listeners with normal hearing and with A-HL did not differ from one another. However, the d-prime values of the listeners with C-HL were significantly higher than that of the other groups. These results indicate that the listeners with C-HL were better able to hear the isolated second formants within the words. The lower d-prime values for the listeners with NH and with A-HL suggest that they incorrectly matched the isolated formants to intact words more often.

The ability of the listeners with C-HL to match the isolated formants to the words from which they originated is more apparent from the results for each talker. Figure 6 shows d-prime and percent correct performance as a function of group. The parameter in this figure is talker. These data show that performance was lower for the child talker relative to the adult talkers. A repeated measures ANOVA confirmed a significant main effect of talker ($F_{2,54} = 29.364$; p<0.000); but no group x talker interaction was revealed ($F_{4,54} = 1.299$; p=0.282). Recall that perceptual coherence was expected to be strongest for the male talker and poorest for the child talker based on the differences in their fundamental frequencies. However, this was not the case. Instead, the perceptual coherence of all three groups was greatest for the child talker.

Finally, several Pearson correlation coefficients were calculated to determine the relation between a listener's ability to match formants to the words from which they originated (d-prime) and his/her age, hearing threshold, and sensation level. Significance was defined as *p*<0.01. Although a significant correlation was observed between age and hearing threshold at 4000 Hz (*r*=0.66 left ear, *r*=0.68 right ear) and at 8000 Hz (*r*=0.66 left ear, *r*=0.60 right ear), no significant correlation was observed between d-prime and age (*r*= -0.30, *p*=0.102). These

results are consistent with decreasing hearing sensitivity with increasing age and confirm that the rather large difference in ages across the groups (nearly 40 years) did not contribute significantly to performance. A significant correlation was also observed between d-prime and hearing threshold in left and right ears at 1000 Hz but the relation was relatively weak ($r = 0.47$ left ear, $r = 0.48$ right ear). These results suggest that the listener's level of performance increased somewhat as hearing loss increased at 1000 Hz. However, because the thresholds at this frequency differed most across groups, the significant correlation likely reflects the overall effect of group rather than hearing sensitivity. Finally, the relation between d-prime and stimulus sensation level at each of the audiometric frequencies (250 through 8000 Hz) was calculated for the listeners with hearing loss only. No significant correlation was found for any frequency in either ear. These results confirm that the difference in d-prime between the listeners with A-HL and C-HL groups was not due to the amplification parameters they received.

C. Task 3: Coherence of amplitude modulated sinusoids

D-prime and maximized performance was also calculated for the amplitude modulated sinusoids. Average maximized performance (bars re: left axis) and d-prime (filled circles re: right axis) are displayed as a function of group in Figure 7. The error bars represent ±1 standard error. On average, all three groups performed at chance levels. A one-way ANOVA of the d-prime data revealed no significant difference between the groups ($F_{2,29} = 1.132$; p=.337). These data suggest strong perceptual coherence for these stimuli in that none of the listeners were able to hear the mid-frequency component in the same manner as the second formant of the speech stimuli.

IV. DISCUSSION

The purpose of this study was to examine the perceptual coherence for both speech and non-speech stimuli in listeners with hearing loss and in listeners with normal hearing. Overall, the results suggested that perceptual coherence differs for speech and non-speech stimuli and is influenced by the onset of hearing loss more so than the hearing loss itself. Specifically, the perceptual coherence for the speech stimuli in listeners with adult-onset hearing losses was equivalent to that of the listeners with normal hearing. These results suggest that perceptual coherence remains even in the presence of a condition that is associated with significant peripheral abnormalities (e.g., wide auditory filters, poor temporal resolution, poor frequency selectivity). In contrast, the listeners with childhood hearing losses showed significantly weaker perceptual coherence for speech (they were better able to hear isolated formants) suggesting that the presence of hearing loss early in life affected their performance. The results for the non-speech stimuli, on the other hand, revealed similarly strong perceptual coherence across groups. That is, all three groups were equally poor at identifying a single acoustic component within a three-tone complex.

These results are difficult to reconcile with those of previous studies for two reasons. First, a consistent theme throughout much of the research in this area is the idiosyncratic performance of listeners with hearing loss (Grose & Hall, 1996; Rose & Moore, 1997). In the present study, the only significant departure from the norm was observed for the listeners with childhood hearing losses. It is worth noting that similar variation in performance would have been observed if the hearing-impaired listeners had been combined into a single group as is often the convention for studies in this population. Second, attempts have been made to relate the performance of listeners with hearing loss to the peripheral abnormalities that accompany cochlear lesions (e.g., wide auditory filters, poor temporal resolution, poor frequency selectivity). However, no direct relationship has been found (Hall & Grose, 1989; Hall & Grose, 1994; Kidd et al., 2002). As for the present study, the results for the listeners with childhood

hearing losses are even more difficult to explain on the basis of peripheral abnormalities. That is, it is unlikely that these abnormalities served to improve their ability to discriminate the individual acoustic components within the speech stimuli. Although one might speculate that the peripheral abnormalities associated with childhood hearing loss differ from those of acquired hearing loss, a direct examination of the psychophysical characteristics of listeners with early vs. late onset hearing loss would provide valuable insight into this issue.

It is also possible that experience with speech perception contributed substantially to each group's performance. Although listeners with adult-onset hearing loss no longer receive a fully intact speech signal, their lengthy experience with speech processing may allow them to compensate for signal degradation, particularly in highly contextual situations. The same may be true for perceptual coherence. As for the listeners with childhood hearing losses, they were expected to demonstrate significantly stronger perceptual coherence based on the assumption that their ability to utilize the subtle acoustic elements of speech is underdeveloped. This hypothesis was motivated by the findings of Nittrouer and Crowther (2001) who reported stronger perceptual coherence in normally hearing children than in adults. For the present study however, the listeners with childhood hearing losses were better able to hear specific acoustic elements within the speech signal than either of their normal-hearing and hearing-impaired counterparts. These results may reflect delayed, arrested, or impaired perceptual development. However, this interpretation should be considered speculative until the characteristics and development of perceptual coherence in children are further defined.

Recall that differences in perceptual coherence were expected across talkers on the basis of fundamental frequency (96, 180, and 227 Hz for the male, female and child talkers, respectively). This expectation was based on the results of Carrell and Opie (1992) who reported significantly improved intelligibility for sine-wave sentences that were amplitude comodulated at low frequencies (50 and 100 Hz) relative to the same sentences comodulated at a higher frequency (200 Hz). This aspect of the study was motivated by a desire to determine whether or not the effects observed using synthetic replicas of speech may be generalized to naturally produced speech. The results of this study suggest that they do not generalize well. By way of example, had fundamental frequency played a substantial role in perceptual coherence, then the results for the female and child talkers should have been more similar given their proximity in frequency (within 50 Hz). Yet the poorest perceptual coherence occurred for the female talker while the strongest perceptual coherence occurred for the child talker. These results suggest that other acoustic/phonetic characteristics likely contribute to perceptual coherence and that fundamental frequency may only play a minor role.

Finally, because little evidence is available regarding the practical implications of perceptual coherence on a listener's ability to perceive speech, the effects of relatively weak or strong coherence can only be speculated. Some have argued that perceptual coherence may assist a listener to perceive speech in noise (Carrell & Opie, 1992; Gordon, 1997; Gordon, 1997). If so, the listener with childhood hearing loss demonstrating weak perceptual coherence may then be expected to have more difficulty perceiving speech in noise than listeners with adult-onset hearing losses. However, the latter group is better known for their complaints regarding difficulty in noise. To complicate the argument, the listeners with adult-onset losses demonstrated perceptual coherence similar to that of the normal-hearing listeners who generally have no complaints regarding noise. Even so, it may be the case that listeners with childhood hearing losses experience more difficulty in noise but they simply have no reference to normal hearing and therefore have no reason to suspect that what they hear is any different than what others hear. We do know that children with even the mildest hearing losses are at risk for poor speech perception, academic performance, and social development in typical classroom noise (Davis et al., 1986; Crandell, 1993; Briscoe et al., 2001). We also know that the vocabulary development of children with hearing loss tends to be delayed relative to that

of children with normal hearing by as much as 2 years, and that the delay persists throughout childhood and increases with increasing severity of hearing loss (Briscoe et al., 2001; Pittman et al., 2005). Unfortunately, there are no data regarding the outcomes of children with hearing loss when they mature and are absorbed into the much larger population of adults with acquired hearing losses. Because the results of the present study suggest that perceptual coherence in adults is detrimentally affected when hearing loss is present early in life, it is possible that other perceptual processes are also affected over the long term. Therefore, it may be wise to examine directly the effects of childhood hearing losses in adults and to consider hearing history as a potential source of variance in all studies involving adults with hearing loss.

ACKNOWLEDGEMENTS

REFERENCES

American National Standards Institute. Methods for calculation of the speech intelligibility index (ANSI S3.5-1997). New York: NY: 1997.

Barker J, Cooke M. Is the sine-wave speech cocktail party worth attending? Speech Communication 1999;27:159–174.

Best CT, Studdert-Kennedy M, Manuel S, Rubin-Spitz J. Discovering phonetic coherence in acoustic patterns. Percept. Psychophys 1989;45:237–250. [PubMed: 2710622]

Bregman, AS. Auditory Scene Analysis. MIT Press; Cambridge, MA: 1990.

Briscoe J, Bishop DV, Norbury CF. Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. Journal of Child Psychology and Psychiatry and Allied Disciplines 2001;42:329–340.

Carrell TD, Opie JM. The effect of amplitude comodulation on auditory object formation in sentence perception. Percept. Psychophys 1992;52:437–445. [PubMed: 1437476]

Crandell CC. Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. Ear and Hearing 1993;14:210–216. [PubMed: 8344478]

Davis JM, Elfenbein J, Schum R, Bentler RA. Effects of mild and moderate hearing impairments on language, educational, and psychosocial behavior of children. Journal of Speech and Hearing Disorders 1986;51:53–62. [PubMed: 3945060]

de Cheveigne A, Kawahara H. YIN, a fundamental frequency estimator for speech and music. J. Acoust. Soc. Am 2002;111:1917–1930. [PubMed: 12002874]

Eisenberg LS, Dirks DD, Bell TS. Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing. Journal of Speech and Hearing Research 1995;38:222–233. [PubMed: 7731213]

Gordon PC. Coherence masking protection in speech sounds: the role of formant synchrony. Percept. Psychophys 1997;59:232–242. [PubMed: 9055618]

Gordon PC. Coherence masking protection in brief noise complexes: effects of temporal patterns. Journal of the Acoustical Society of America 1997;102:2276–2283. [PubMed: 9348685]

Grose JH, Hall JW. Comodulation masking release: is comodulation sufficient? Journal of the Acoustical Society of America 1993;93:2896–2902. [PubMed: 8315153]

Grose JH, Hall JW. Cochlear hearing loss and the processing of modulation: effects of temporal asynchrony. Journal of the Acoustical Society of America 1996;100:519–527. [PubMed: 8675845]

Hall JW, Grose JH. Spectrotemporal analysis and cochlear hearing impairment: effects of frequency selectivity, temporal resolution, signal frequency, and rate of modulation. Journal of the Acoustical Society of America 1989;85:2550–2562. [PubMed: 2745879]

Hall JW, Grose JH. Signal detection in complex comodulated backgrounds by normal-hearing and cochlear-impaired listeners. Journal of the Acoustical Society of America 1994;95:435–443. [PubMed: 8120254]

Hall JW, Haggard MP, Fernandes MA. Detection in noise by spectro-temporal pattern analysis. Journal of the Acoustical Society of America 1984;76:50–56. [PubMed: 6747111]

Kidd G, Arbogast TL, Mason CR, Walsh M. Informational masking in listeners with sensorineural hearing loss. J. Assoc. Res. Otolaryngol 2002;3:107–119. [PubMed: 12162362]

Marshall L, Jesteadt W. Comparison of pure-tone audibility thresholds obtained with audiological and two-interval forced-choice procedures. Journal of Speech and Hearing Research 1986;29:82–91. [PubMed: 3702383]

McFadden D. Comodulation detection differences using noise-band signals. Journal of the Acoustical Society of America 1987;81:1519–1527. [PubMed: 3584689]

Nittrouer S, Crowther CS. Coherence in children's speech perception. Journal of the Acoustical Society of America 2001;110:2129–2140. [PubMed: 11681390]

Pittman AL, Lewis DE, Hoover BM, Stelmachowicz PG. Rapid word-learning in normal-hearing and hearing-impaired children: effects of age, receptive vocabulary, and high-frequency amplification. Ear and Hearing 2005;26:619–629. [PubMed: 16377997]

Remez R, Rubin PE. On the perception of speech from time-varying acoustic information: contributions of amplitude variation. Percept. Psychophys 1990;48:313–325. [PubMed: 2243755]

Remez RE, Pardo JS, Piorkowski RL, Rubin PE. On the bistability of sine wave analogues of speech. Psychol. Sci 2001;12:24–29. [PubMed: 11294224]

Rose MM, Moore BC. Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. J. Acoust. Soc. Am 1997;102:1768–1778. [PubMed: 9301054]

Seewald, RC.; Cornelisse, LE.; Ramji, KV.; Sinclair, ST.; Moodie, KS.; Jamieson, DG. DSL v4.1 for Windows: A software implementation of the Desired Sensation Level (DSL[i/o]) Method for fitting linear gain and wide-dynamic-range compression hearing intruments. User's manual. 1997.
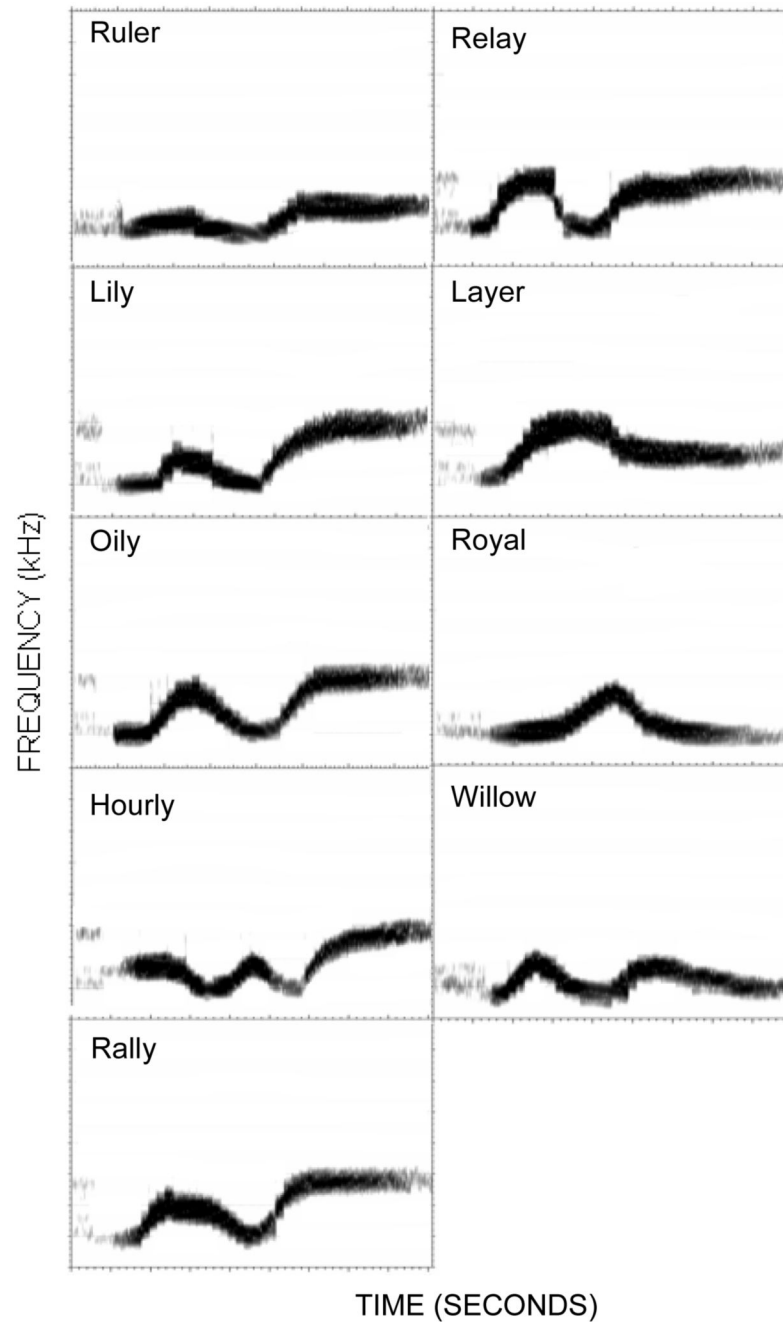
1.
Spectrograms of the second formants for the female talker. The abscissa and ordinate of each spectrogram extend to 8000 Hz and 1000 ms, respectively.
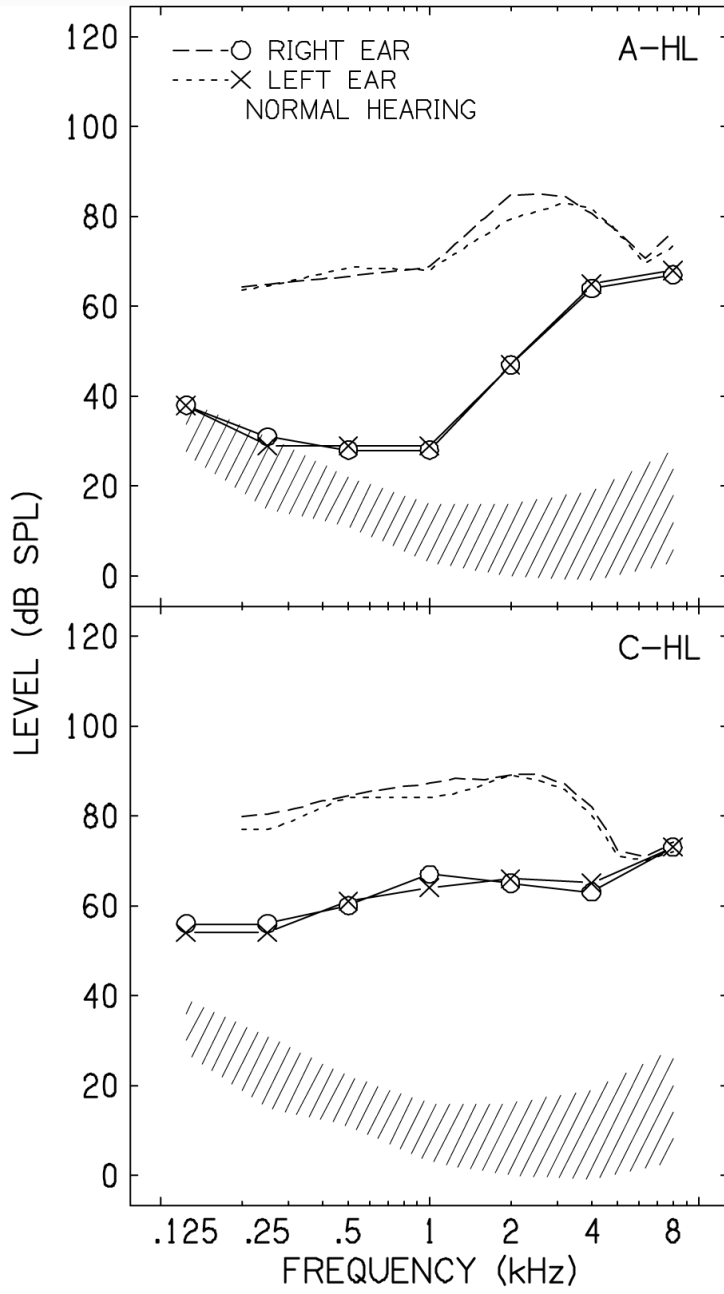
2.
Long-term average speech spectra (dashed lines) and hearing thresholds (solid lines) for the left and right ears of the listeners with adult onset hearing loss (A-HL) in the upper panel and with longstanding, childhood hearing loss (C-HL) in the lower panel. The hatched area in each panel represented the range of hearing thresholds of the listeners with normal hearing as defined by one standard deviation around the mean.
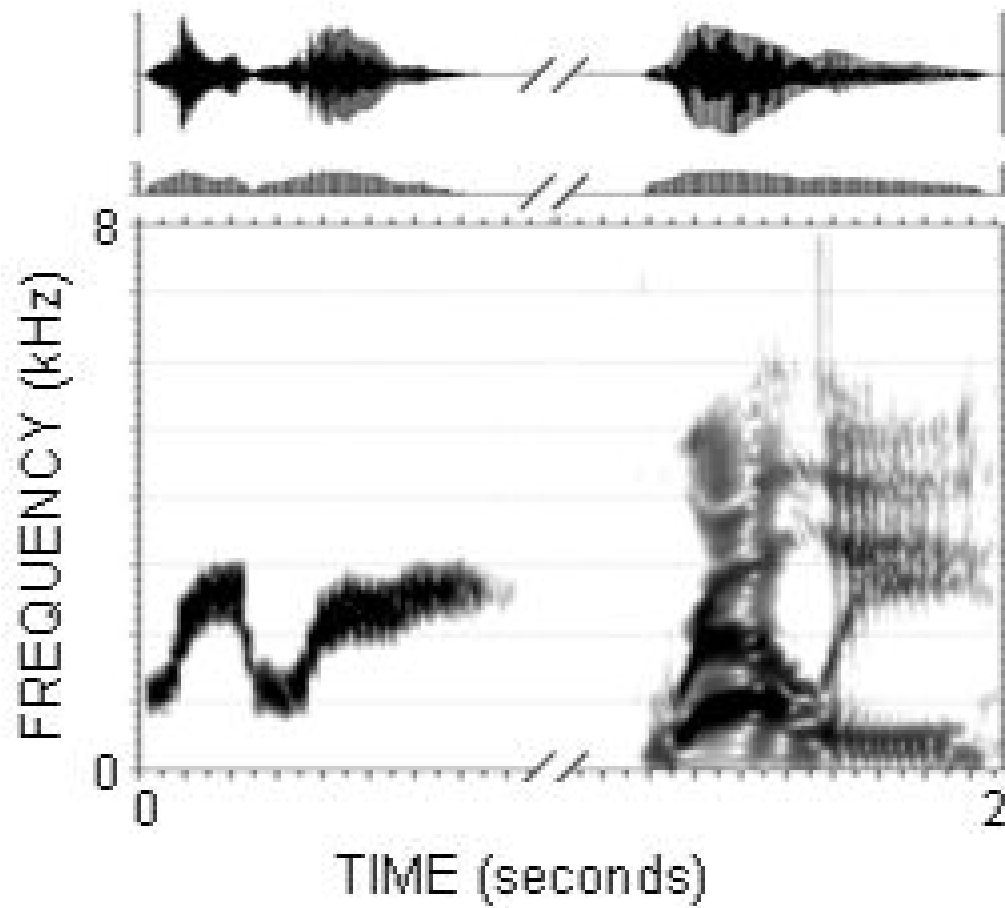
3.

An example of a stimulus set used in the primary task. A second formant was paired with an intact, neighboring word separated by 800 ms. In this example of a 'no' trial, the second formant *relay* was paired with word *rally*.
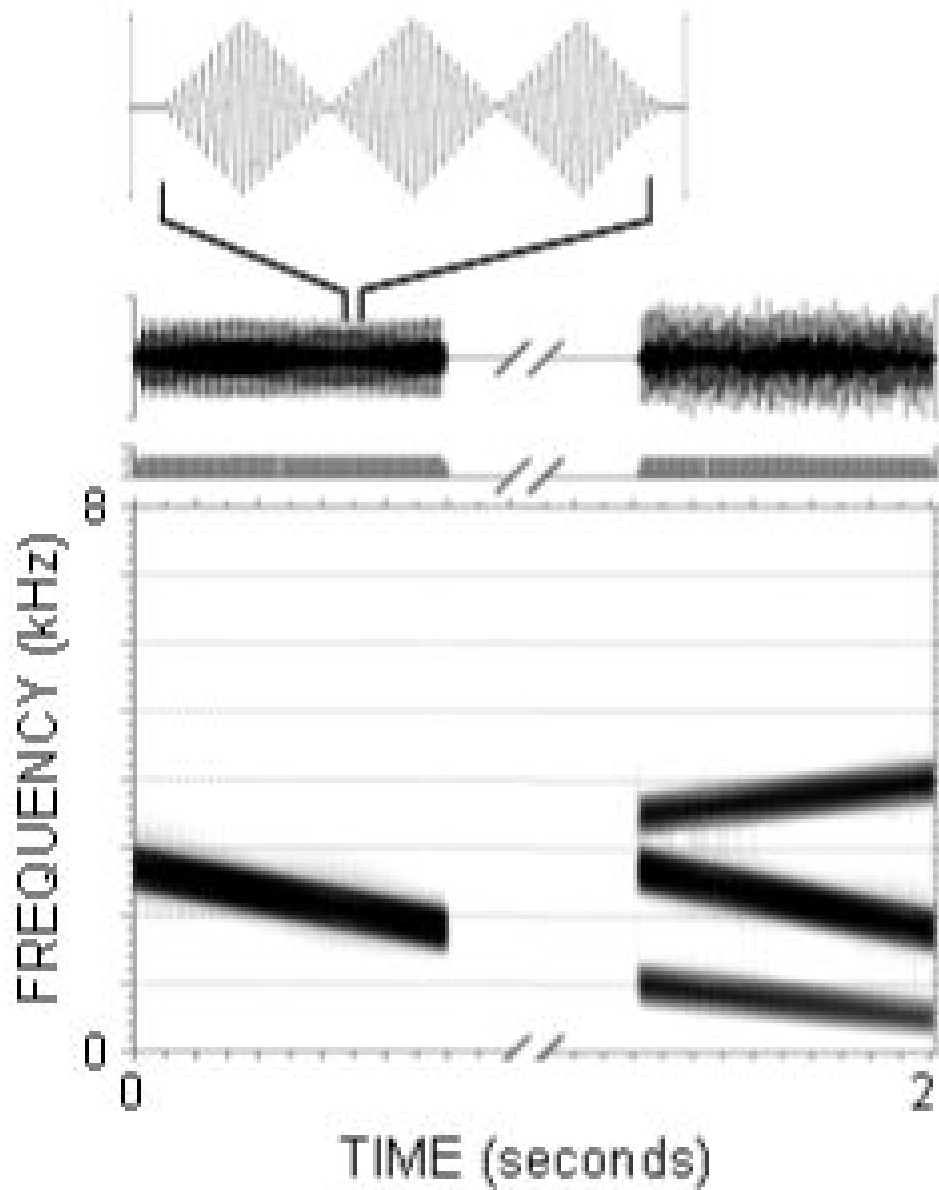
4.
An example of a stimulus set used in the final task. A mid-frequency amplitude modulated sinusoid was paired with a 3-tone complex. In this example of a 'yes' trial, the mid-frequency sinusoid was contained within the 3-tone complex. The upper waveform shows three modulation cycles of the mid-frequency sinusoid.
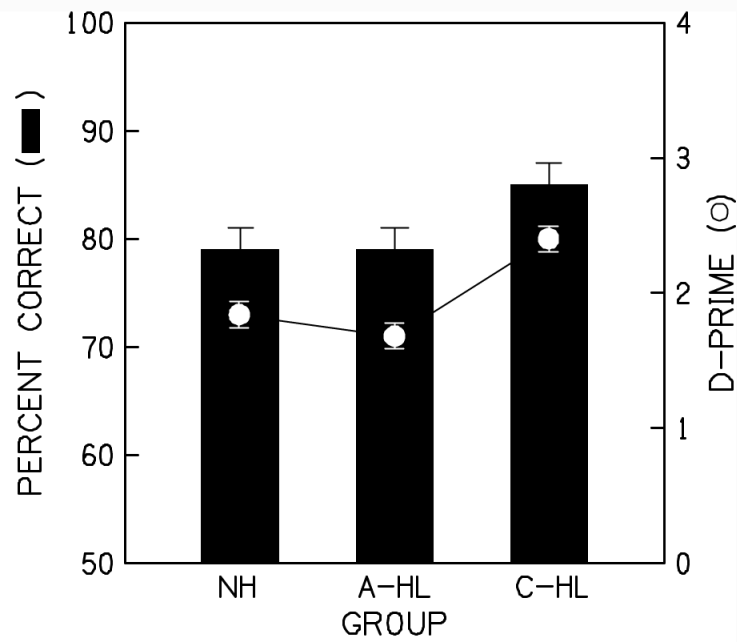
5.
Average maximized performance and d-prime values for the speech task as a function of group. Performance (bars) is referenced to the left axis and d-prime (circles) is referenced to the right axis. All error bars are +1 SE. Groups are listeners with normal hearing (NH), adult-onset hearing loss (A-HL), and childhood hearing loss (C-HL).
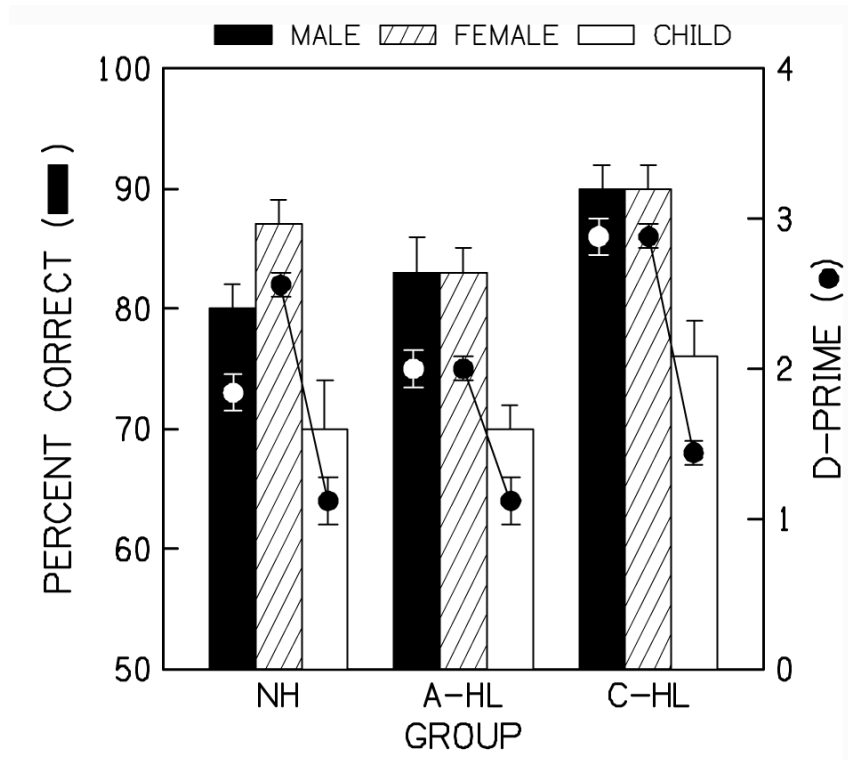
6.
Same convention as in Figure 5 but with the parameter of talker. Filled, hatched, and open bars represent the male, female, and child talkers, respectively.
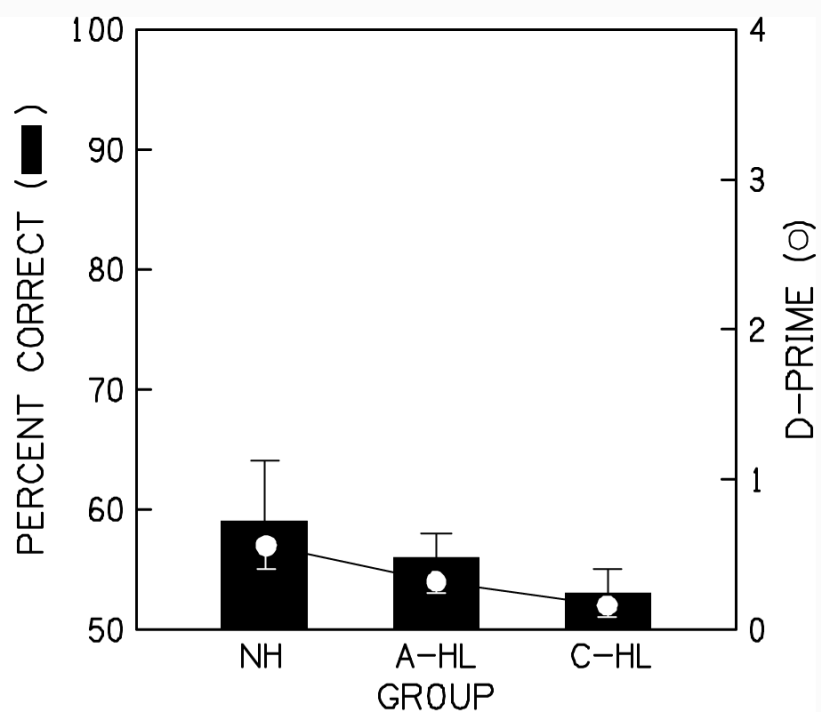
7.
Same convention as Figure 5 but for the amplitude modulated sinusoid stimuli.

Table I

Age and hearing thresholds for the listeners with adult-onset hearing losses (A–HL) and childhood hearing losses (C-HL)

ID #	Age	Ear	Hearing Thresholds (dB SPL)						
			0.125	0.25	0.5	1	2	4	8
A-HL1	67	L	27	23	15	12	60	65	67
		R	32	33	20	13	67	60	58
A-HL2	60	L	43	38	38	33	67	70	73
		R	35	37	33	43	58	73	87
A-HL3	57	L	33	17	23	18	37	63	43
		R	33	17	23	23	48	70	47
A-HL4	69	L	27	5	7	7	15	60	38
		R	17	15	15	8	37	67	50
A-HL5	68	L	30	22	33	30	30	60	73
		R	30	13	13	13	35	60	73
A-HL6	66	L	33	22	13	13	57	65	77
		R	27	23	13	23	67	67	77
A-HL7	66	L	30	17	17	20	15	83	87
		R	32	26	17	8	15	63	70
A-HL8	65	L	37	33	27	37	67	67	93
		R	52	43	30	37	53	63	77
A-HL9	58	L	57	66	61	69	73	65	46
		R	69	54	56	61	45	57	58
A-HL10	64	L	58	50	57	53	46	51	82
		R	48	48	58	50	48	55	68
AVG (1 SE)	64	L	38 (3)	29 (4)	29 (4)	29 (4)	47 (5)	65 (3)	68 (4)
		R	38 (4)	31 (4)	28 (4)	28 (4)	47 (4)	64 (2)	67 (4)
C-HL11	39	L	55	53	53	53	63	53	83
		R	58	58	58	58	58	58	83
C-HL12	35	L	73	83	93	92	83	70	68
		R	83	78	88	93	87	68	68
C-HL13	19	L	53	37	38	38	23	13	33
		R	33	33	38	47	38	5	13
C-HL14	14	L	61	63	71	68	63	63	59
		R	67	73	68	68	63	63	62
C-HL15	54	L	63	63	78	78	83	73	88
		R	53	43	53	68	68	68	78
C-HL16	35	L	66	72	76	74	68	67	81
		R	62	63	63	68	67	62	73
C-HL17	36	L	38	43	53	78	79	66	73
		R	43	43	56	88	83	71	75
C-HL18	41	L	36	33	33	37	43	57	33
		R	43	38	36	28	33	48	36
C-HL19	44	L	57	48	51	58	78	108	130
		R	78	81	78	77	68	91	107
C-HL20	43	L	35	48	62	68	78	83	78
		R	43	48	63	73	83	98	130
AVG (1 SE)	36	L	54 (4)	54 (4)	61 (4)	64 (4)	66 (4)	65 (5)	73 (5)
		R	56 (4)	56 (4)	60 (4)	67 (4)	65 (4)	63 (5)	73 (6)
Normal Hearing AVG (1 SE)	25	L	33 (2)	21 (3)	17 (2)	8 (3)	6 (3)	11 (3)	14 (4)
		R	34 (3)	25 (2)	16 (2)	12 (2)	10 (2)	8 (3)	17 (3)

Table II

On- and off-set frequency of each component in the time-varying sinusoid stimuli

Stim #	Component	Onset (Hz)	Offset (Hz)
1	low	1000	500
	mid	3000	1500
	high	3500	4000
2	low	1000	500
	mid	2800	1800
	high	3500	4000
3	low	1000	500
	mid	2500	2000
	high	3500	4000
4	low	1000	500
	mid	2000	2500
	high	3500	4000
5	low	1000	500
	mid	1800	2800
	high	3500	4000
6	low	1000	500
	mid	1500	3000
	high	3500	4000

Table III

Pairing of the isolated second formants with the intact words. "Yes" trials contained a second formant and the intact word from which it originated. "No" trials contained a second formant and an intact word neighboring the word from which the second formant was obtained

		Intact Words								
		ruler	lily	oily	hourly	rally	relay	layer	royal	willow
2nd Formant	lily	No	Yes	No						
	oily		No	Yes	No					
	hourly			No	Yes	No				
	rally				No	Yes	No			
	relay					No	Yes	No		
	layer						No	Yes	No	
	royal							No	Yes	No

Table IV

Pairing of the mid-frequency components with the three-tone complexes. "Yes" trials contained a mid-frequency component and the three-tone complex from which it originated. "No" trials contained a mid-frequency component and a neighboring three-tone complex

		Three-Tone Complex					
		1	2	3	4	5	6
Mid-Frequency Component	2	No	Yes	No			
	3		No	Yes	No		
	4			No	Yes	No	
	5			No	No	Yes	No