Effect of age, presentation method, and learning on identification of noise-vocoded words

Signy Sheldon, M. Kathleen Pichora-Fuller, and Bruce A. Schneider

Citation: The Journal of the Acoustical Society of America **123**, 476 (2008); doi: 10.1121/1.2805676
View online: https://doi.org/10.1121/1.2805676
View Table of Contents: https://asa.scitation.org/toc/jas/123/1
Published by the Acoustical Society of America

ARTICLES YOU MAY BE INTERESTED IN

Priming and sentence context support listening to noise-vocoded speech by younger and older adults
The Journal of the Acoustical Society of America **123**, 489 (2008); https://doi.org/10.1121/1.2783762

Effectiveness of computer-based auditory training in improving the perception of noise-vocoded speech
The Journal of the Acoustical Society of America **121**, 2923 (2007); https://doi.org/10.1121/1.2713668

Rapid perceptual learning of noise-vocoded speech requires attention
The Journal of the Acoustical Society of America **131**, EL236 (2012); https://doi.org/10.1121/1.3685511

How young and old adults listen to and remember speech in noise
The Journal of the Acoustical Society of America **97**, 593 (1995); https://doi.org/10.1121/1.412282

Effects of envelope bandwidth on the intelligibility of sine- and noise-vocoded speech
The Journal of the Acoustical Society of America **126**, 792 (2009); https://doi.org/10.1121/1.3158835

Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability
The Journal of the Acoustical Society of America **61**, 1337 (1977); https://doi.org/10.1121/1.381436

ASA

Advance your science and career
as a member of the
ACOUSTICAL SOCIETY OF AMERICA

LEARN MORE

Effect of age, presentation method, and learning on identification of noise-vocoded words

Signy Sheldon, M. Kathleen Pichora-Fuller,[a)] and Bruce A. Schneider
Department of Psychology, University of Toronto, 3359 Mississauga Road North, Mississauga, Ontario L5L 1C6, Canada

(Received 31 October 2006; revised 2 October 2007; accepted 15 October 2007)

Noise vocoding was used to investigate the ability of younger and older adults with normal audiometric thresholds in the speech range to use amplitude envelope cues to identify words. In Experiment 1, four 50-word lists were tested, with each word presented initially with one frequency band and the number of bands being incremented until it was correctly identified by the listener. Both age groups required an average of 5.25 bands for 50% correct word identification and performance improved across the four lists. In Experiment 2, the same participants who completed Experiment 1 identified words in four blocked noise-vocoded conditions (16, 8, 4, 2 bands). Compared to Experiment 1, both age groups required more bands to reach the 50% correct word identification threshold in Experiment 2, 6.13, and 8.55 bands, respectively, with younger adults outperforming older adults. Experiment 3 was identical to Experiment 2 except the participants had no prior experience with noise-vocoded speech. Again, younger adults outperformed older adults, with thresholds of 6.67 and 8.97 bands, respectively. The finding of age effects in Experiments 2 and 3, but not in Experiment 1, seems more likely to be related to differences in the presentation methods than to experience with noise vocoding. © *2008 Acoustical Society of America.*
[DOI: 10.1121/1.2805676]

PACS number(s): 43.71.Lz, 43.66.Sr, 43.71.Gv [MSS] Pages: 476–488

I. INTRODUCTION

Older adults often report more difficulties than younger adults in understanding spoken language, especially in adverse listening conditions (for reviews see CHABA, 1988; Divenyi and Simon, 1999; Pichora-Fuller and Souza, 2003). Even older adults with normal audiometric thresholds in the speech range have these difficulties, so it seems unlikely that loss of audibility can fully explain their poor comprehension. Given that auditory temporal processing is an integral part of spoken language comprehension (e.g., de Boer and Dreschler, 1987; van Tasell *et al.*, 1987; Rosen, 1992; Shannon *et al.*, 1995), many have suggested that age-related declines in auditory temporal processing may contribute to the comprehension difficulties that older adults often have in challenging conditions (for reviews see Fitzgibbons and Gordon-Salant, 1996; Schneider and Pichora-Fuller, 2001; Versfeld and Dreschler, 2002; Pichora-Fuller and Souza, 2003).

A. Effect of age on speech and temporal processing

Temporally coded cues relevant to speech processing have been described at three levels: subsegmental (voice), segmental (phonemic), and suprasegmental (syllabic and lexico-syntactic). Subsegmental fine structure cues include periodicity cues based on the fundamental frequency and harmonic structure of the voice. Segmental information is provided by local gap and duration cues in the envelope which contribute to phoneme identification (e.g., presence or absence of a stop consonant, voicing). Suprasegmental cues, such as amplitude fluctuations in the region of 3–20 Hz, convey prosodic information involving the rate and rhythm of speech and are used in lexical and syntactic processing (Rosen, 1992; Philips, 1995; Greenberg, 1996; Schneider and Pichora-Fuller, 2001; Shannon, 2002; Pichora-Fuller and Souza, 2003). It is important to know whether age affects the temporal processing of speech cues at one or more of these three levels within and across spectral regions [for a discussion see also Souza and Boike (2006)]. While there is strong evidence of age effects on aspects of auditory temporal processing that are relevant to speech processing at the subsegmental and segmental levels, less is known about how age affects the processing of suprasegmental speech cues.

With respect to the subsegmental level of temporal processing, physiological and behavioral studies suggest that there are age-related decrements in synchrony coding at various stages of auditory processing which could undermine the periodicity coding of speech and nonspeech signals (e.g., Frisina, 2001; Pichora-Fuller *et al.*, 2007). For example, the pattern of binaural masking-level differences in younger and older adults suggests that the precision of periodicity coding is reduced with age (Pichora-Fuller and Schneider, 1992). Monaurally, frequency difference limens at lower frequencies are larger for older than for younger adults, but this age-related difference is less marked at higher frequencies where periodicity coding does not play as significant a role (Abel *et al.*, 1990). Furthermore, older adults have more difficulty discriminating a mistuned harmonic in a harmonic complex, especially for short duration sounds (Alain *et al.*, 2001), and older adults have more problems than younger

[a)]Electronic mail: k.pichora.fuller@utoronto.ca

 © 2008 Acoustical Society of America

adults in segregating concurrent vowels (Summers and Leek, 1998; Vongpaisal and Pichora-Fuller, 2007). These studies demonstrate the mounting evidence pointing to age-related deficits in the temporal processing of subsegmental or fine-structure cues that are believed to play a role in voice identification and segregation (e.g., de Cheveigné, 2003).

Other studies have demonstrated age-related differences in temporal processing relevant to the segmental level. The effects of age are seen clearly in a large number of gap detection experiments in which listeners are asked to detect the presence of a gap between sound markers. In general, older adults do not detect a gap until it is significantly longer than the smallest gap that can be detected by younger adults for either nonspeech or speech markers (e.g., Snell, 1996; Pichora-Fuller *et al.*, 2006). For example, the thresholds of older adults were approximately twice as large as those of younger adults when detecting a gap between two Gaussian-enveloped tone pips (Schneider *et al.*, 1994). Similarly, mean gap thresholds were significantly larger for older listeners compared to younger listeners for low-passed filtered noise bursts (Snell and Frisina, 2000). Age-related deficits in temporal processing of segmental-level speech cues may also arise from declines in duration discrimination. In duration discrimination studies, younger and older listeners are asked to identify the longer or shorter of two stimuli. Older adults have more difficulty with this task than younger adults for both nonspeech and speech stimuli (e.g., Abel *et al.*, 1990; Bergerson *et al.*, 2001; Fitzgibbons and Gordon-Salant, 1994, 1995; Gordon-Salant *et al.*, 2006).

Suprasegmental cues involving variations in pitch, loudness, and/or timing contribute to speech prosody (e.g., Cutler *et al.*, 1997). Variations in timing include changes in the duration of phonemes and words to alter the rate and rhythm of speech. Previous research provides evidence that older adults may be more disadvantaged than younger adults when prosodic cues are disrupted by various types of temporal distortion, but that they may benefit as much or more than younger adults when prosodic cues are available, especially in challenging listening conditions.

On the one hand, various methods of temporally distorting speech, including speeding or time compression, have a more deleterious effect on the spoken language understanding of older compared to younger listeners on word tests (e.g., Sticht and Gray, 1969; Konkle *et al.*, 1977; Stuart and Phillips, 1996), sentence tests (e.g., Gordon-Salant and Fitzgibbons, 1993, 1997), and discourse tests (e.g., Vaughan and Letowski, 1997; for a review see Wingfield, 1996). In addition to the possible cognitive strain introduced when speech is speeded, speeding speech may also have negative consequences on the comprehension abilities of older adults because their auditory systems are more susceptible than are those of younger adults to the acoustical temporal distortions that some methods introduce in the speech signal (e.g., Wingfield *et al.*, 1999; Schneider *et al.*, 2005). Since time compression alters the envelope of speech, it may be that older adults are more affected than younger adults by distortions in the shape of the envelope that might compromise the auditory processing of cues used at the phonemic, lexical, and/or syntactic levels.

On the other hand, evidence that the use of envelope cues is preserved in older adults comes from studies showing that younger and older adults both benefit from duration and envelope cues to identify words (e.g., Wingfield *et al.*, 2000), and that both age groups benefit from the insertion of pauses to understand sentences (e.g., Wingfield *et al.*, 1999). There is also evidence that older adults benefit more than younger adults from prosodic cuing to understand speeded sentences (e.g., Wingfield *et al.*, 1992), or short passages (e.g., Stine and Wingfield, 1987). Furthermore, in conversational discourse older adults demonstrate preserved use of prosody to understand socio-emotional relational information even when hearing loss impedes the understanding of content because of poor ability to identify phonemes and words (Villaume *et al.*, 1994).

Since temporal amplitude envelope cues are among the suprasegmental cues that contribute to speech prosody and such cues can be used even by those with significant hearing loss (Turner *et al.*, 1995), it seems likely that the findings of preserved ability in older adults to use prosody and their greater reliance on it in challenging conditions may be, at least partially, attributable to their good ability to use envelope information when they are listening to words, sentences, or discourse. Souza and Boike (2006) suggest a disassociation between auditory cue use and age—younger listeners tend to use both temporal envelope and fine structure cues, whereas older listeners seem to rely more heavily on suprasegmental envelope cues than on spectral or temporal fine structure cues. Thus, one hypothesis is that use of these suprasegmental cues does not decline with age, but rather that the use of suprasegmental cues may even compensate for the reduced ability of older listeners to use subsegmental and segmental speech cues.

The divergent theories concerning the abilities of older adults to use temporal amplitude envelope cues make it an issue that warrants further investigation. Therefore, the primary goal of this study is to examine the effect of age on the ability of listeners to use envelope cues when fine structure cues are minimized. To this end, we tested word identification in younger and older listeners using noise-vocoded speech.

B. Noise vocoding

Noise vocoding is a form of speech distortion that involves dividing a speech signal into specific frequency bands, and then, within each band, extracting the temporal amplitude envelope and using it to modulate noise of the same bandwidth. In effect, the fine structure of the signal is replaced with noise. Thus, noise vocoding preserves temporal amplitude envelope cues within specific frequency bands and eliminates fine structure cues, including periodicity cues (see Fig. 1). As the number of bands is increased, more band-specific envelope information becomes available. It has been shown that the intelligibility of noise-vocoded speech stimuli is dependent on the number of frequency bands used in voc-

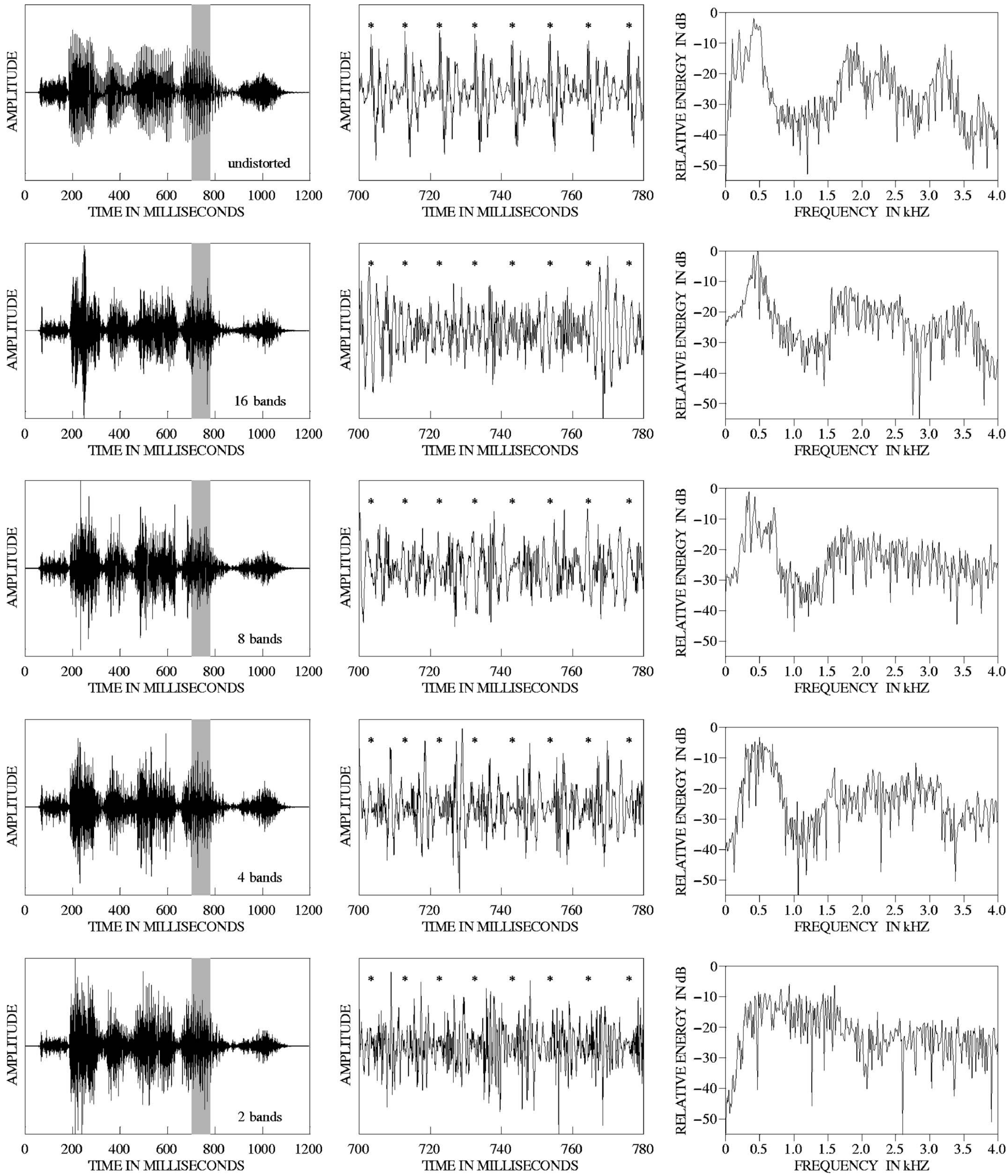

FIG. 1. Left panels, from top to bottom: The time wave forms of the undistorted followed by the vocoded versions (16, 8, 4, and 2 bands, respectively) of the sentence "Say the word ace." The shaded portion identifies an 80-ms segment of the vowel in the word ace. The middle panel, from top to bottom, shows the time wave forms of the undistorted (top) followed by the vocoded versions of the 80-ms segment of the vowel. The periodicity of the vowel in the undistorted case is indicated by the asterisks in the top middle panel. Note that the periodicity is virtually eliminated during vocoding. Right panels, from top to bottom: The energy spectra for the undistorted (top) followed by the vocoded versions of the 80-ms segment of the vowel.

oding (e.g., Loizou *et al.*, 1999; Shannon *et al.*, 1995). In a pivotal study by Shannon and colleagues (1995), near perfect levels of speech recognition were achieved with as few as four frequency bands for young adult listeners with good hearing. Not only did this study provide convincing evidence of the importance of envelope cues for spoken language

TABLE I. Mean (s.d.) of audiometric air-conducted pure-tone thresholds (dB HL) of the test ears for the younger and older participants in Experiments 1, 2, and 3.

	Frequency (kHz)							
	0.25	0.5	1	2	3	4	6	8
Younger participants (Experiments 1 and 2; $N=12$)								
Mean dB HL	6.25	4.58	1.67	0.833	0.833	1.67	7.98	5.42
s.d.	(4.33)	(8.11)	(3.89)	(3.60)	(5.57)	(6.51)	(7.82)	(6.90)
Younger participants (Experiment 3; $N=12$)								
Mean dB HL	6.25	4.17	0.00	3.33	−0.42	−0.42	3.75	0.42
s.d.	(6.78)	(5.57)	(5.22)	(4.92)	(5.82)	(6.20)	(7.42)	(7.21)
Older participants (Experiments 1 and 2; $N=12$)								
Mean dB HL	8.33	5.42	6.25	11.67	10.83	18.75	27.08	27.50
s.d.	(6.62)	(4.96)	(2.61)	(8.91)	(7.06)	(10.03)	(14.37)	(19.48)
Older participants (Experiment 3; $N=12$)								
Mean dB HL	8.33	5.42	6.25	11.67	10.83	22.50	31.25	45.42
s.d.	(6.62)	(4.96)	(2.61)	(8.91)	(7.06)	(9.68)	(14.16)	(19.63)

comprehension, but it also provided an important new method for studying the contribution of envelope cues to speech processing.

C. Perceptual learning

Noise-vocoded speech is not experienced outside of the lab; therefore, it is important to determine how perceptual learning might influence age-related differences in performance. The effect of perceptual learning on speech perception has been shown in behavioral (e.g., Davis *et al.*, 2005) and physiological studies (e.g., Tremblay *et al.*, 1997, 2001; Callan *et al.*, 2003). Furthermore, the effect of perceptual learning is seen for various forms of distorted speech, such as synthetic speech (Greenspan *et al.*, 1988), time-compressed speech (Dupoux and Green, 1997; Peelle and Wingfield, 2005), and noise-vocoded speech (Davis *et al.*, 2005). Specific to noise-vocoded speech, a study by Davis and colleagues (2005) showed that noise-vocoded sentences that were initially unintelligible to participants became markedly more intelligible as listeners gained experience with noise vocoding, implicating the role of perceptual learning in listeners' understanding of noise-vocoded speech.

Few studies have examined how younger and older listeners differ in terms of auditory learning. Among these studies, Peelle and Wingfield (2005) found that perceptual learning for noise-vocoded paragraphs was comparable for older and younger adults. However, in a subsequent experiment, they found an age-related deficit in transferring perceptual learning from one stimulus set to another when the speech was time compressed. Specifically, older adults were worse than younger adults in adapting to one rate of time-compressed speech after having adapted to another speech rate. Therefore, a secondary goal of our study was to investigate how age may affect perceptual learning as listeners acquire and consolidate learning gained from listening to noise-vocoded speech.

II. EXPERIMENT 1

The first goal of Experiment 1 was to determine the number of bands required for younger and older adults to correctly identify noise-vocoded words in a carrier phrase. The second goal was to determine whether or not participants in the two age groups showed improvement in word identification performance as they gained experience listening to noise-vocoded speech.

A. Method

1. Participants

Twelve younger adults (mean age=22.3 years, s.d.=2.2, range=19–25) and twelve older adults (mean age =70.2 years, s.d.=3.1, range=66–74) participated in the experiment. All participants in both age groups had pure-tone air-conduction thresholds in the test ear less than or equal to 25 dB HL from 0.25 to 3 kHz, i.e., they had audiometric thresholds that were clinically normal in the speech range (Table I). All participants had learned English before the age of 5 years and had been educated in English in a country where it is the dominant language. To measure verbal knowledge, each participant was given the Mill-Hill Vocabulary Scale (Raven, 1965). On average, the older group had better vocabulary scores than the younger group [mean for the younger group=12.5/20 and s.d.=2.0; mean for the older group=15.3/20 and s.d.=2.9; $t(22)=-2.68$, $p<0.05$]. On average, the younger group had 1.8 years more education than the older group [mean for the younger group =15.4 years of education and s.d.=1.6; mean for the older group=13.6, s.d.=2.9; $t(22)=1.84$; $p<0.05$]. All of the participants were paid volunteers recruited from the local community who gave informed consent in compliance with the protocol approved by the university's ethics review board. No participant had previously heard noise-vocoded speech.

2. Stimuli and apparatus

The test stimuli were the digital recordings of the four Northwestern University Auditory Test Number 6 (NU-6) word lists that have been standardized for use in clinical speech audiometry (see Penrod, 1985) and distributed on compact disk by Auditec of St. Louis. Each list consists of 50 monosyllabic words spoken by a male talker preceded by the carrier phrase "*Say the word*…." For each word, 16 different

TABLE II. Boundary frequencies for the 1 to 16 band-processed noise-vocoding.

1 band	300	6000															
2 band	300	1528	6000														
3 band	300	814	1528	6000													
4 band	300	722	1528	3066	6000												
5 band	300	546	994	1528	3296	6000											
6 band	300	494	814	1528	2210	3642	6000										
7 band	300	460	706	1083	1528	2549	3911	6000									
8 band	300	477	722	1061	1528	2174	3066	4298	6000								
9 band	300	418	584	814	1136	1528	2210	3083	4301	6000							
10 band	300	405	546	737	994	1528	1810	2443	3296	4447	6000						
11 band	300	394	517	679	892	1171	1528	2019	2650	3480	4570	6000					
12 band	300	385	494	634	814	1045	1528	1722	2210	2837	3642	4674	6000				
13 band	300	378	476	599	754	950	1196	1528	1896	2387	3005	3784	4765	6000			
14 band	300	372	460	570	706	875	1083	1342	1528	2058	2549	3158	3911	4844	6000		
15 band	300	366	447	546	667	814	994	1214	1528	1810	2210	2699	3296	4024	4914	6000	
16 band	300	382	477	590	722	878	1061	1276	1528	1825	2174	2584	3066	3632	4298	5080	6000

noise-vocoded band conditions were created, beginning with a one-band condition and increasing the number of bands, by one, up to a 16-band condition. That is, for every word, there were 17 files: one file for the intact condition and one for each of the 16 conditions differing in the number of bands used during noise vocoding.

To create noise-vocoded stimuli, we followed the procedure described in detail by Eisenberg *et al.* (2000). First, the speech stimuli were converted with the Goldwave digital audio editor to binary files with a sampling rate of 20 kHz. Using MATLAB software, stimuli were processed through a pre-emphasis filter (a high-pass first-order Butterworth infinite impulse response (IIR) filter with a cut-off frequency of 1.2 kHz and a roll-off of 6 dB/octave). The signal was split into a varying number of frequency bands ($n=1-16$) using fourth-order elliptical IIR bandpass filters with a maximum peak-to-peak ripple of 0.5 dB in the passband and a minimum attenuation of 40 dB in the stop band. The passband used to split the signal into frequency bands spanned a frequency range from 0.3 to 6 kHz for all conditions. The frequency spacing of the filter banks was based on the work of Greenwood (1990). The boundary frequencies for the band-processed conditions are shown in Table II. To extract the envelopes, the magnitude of the Hilbert transform was computed and passed through a low-pass filter (second-order Butterworth IIR with cut-off frequency of 0.1 kHz). One difference in procedures was that whereas Eisenberg *et al.* (2000) rectified and then low-pass filtered the filter bank outputs, we extracted the envelope using the magnitude of the Hilbert transform followed by a low-pass filter similar to that used by Eisenberg *et al.* (2000). Narrow-band noise was generated by passing a Gaussian white noise signal through the same Butterworth and elliptical filters. The envelopes extracted in the previous step were then used to modulate the corresponding band of noise. The bands of modulated noise were then summed together. Finally, the stimuli were converted to wav format with a sampling rate of 24.414 kHz using the Goldwave digital audio editor.

During the testing session, the participant was seated comfortably inside an Industrial Acoustics Company double-walled sound-attenuating booth. The stimulus files were played using a Tucker Davis Technologies System III and were presented monaurally to the participant's better ear over a Sennheiser HD 265 headphone. All stimuli were presented at 70 dB SPL for both age groups.

3. Procedure

The procedure for Experiment 1 was adapted from the gating paradigm (Grosjean, 1996). In the original gating procedure, the word is broken into a sequence of time gates. During the first trial, only one gate is presented and the listener attempts to identify the word. On each subsequent trial the number of gates is increased until the listener reliably identifies the word. Thus, the gating procedure involves multiple presentations of portions of the same stimulus with gradual increments in the amount of signal delivered in order to determine how much of the signal must be heard for the word to be correctly identified. The gating procedure was adapted for the present study such that, rather than incrementing the number of time gates presented to the listener, we incremented the number of bands that were presented. Specifically, a word was first presented in the one-band noise-vocoded condition. The participant was asked to identify the word or to respond "I don't know." If the listener did not correctly identify the word, it was presented again in the 2-band condition. The number of bands continued to be incremented by one until the word was correctly identified. If the word was not correctly identified by the 16-band condition, then the word was presented in the intact condition. Feedback was given via a computer monitor. Guessing was strongly encouraged.

Every participant heard every word in each of the four lists; however, individual participants heard a varying number of band conditions for each word depending on the band condition in which he or she correctly identified the word. The words were presented in random order within each list and the list order was counterbalanced across participants so that each list was presented an equal number of times in each order position for each age group.

Each participant completed the four word lists in a single session, with each word list taking approximately

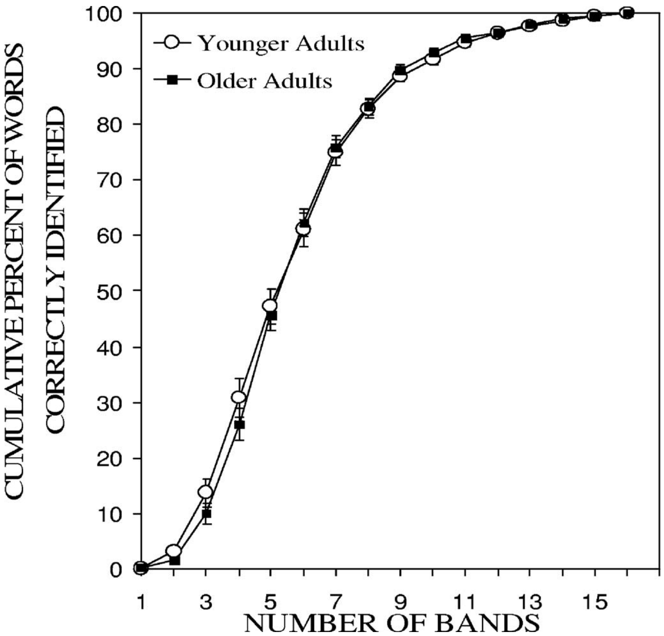

FIG. 2. The cumulative percentage of words correctly identified, averaged across participants, as a function of the number of bands for younger (open circles) and older (closed squares) participants in Experiment 1. Standard error bars are shown.

25 min to complete. Participants were given a 10-min break between each list. The experimenter, who was outside the booth, listened to and immediately scored each response. If the experimenter was uncertain about any response, the participant was asked to repeat and spell the word aloud. Any difference between response and target was marked as an error. All sessions were audiotaped to enable subsequent verification of the scoring.

B. Results and discussion

Words that were only correctly identified in the intact condition or never correctly identified were excluded from the analyses. In total, only 178 out of 4800 responses ($<4\%$) were excluded, of which 70 responses were from younger participants and 108 were from older participants. The cumulative percentage of words identified correctly in each band condition was calculated for each participant for each list. The mean of these values for each group is plotted in Fig. 2, illustrating that word identification performance was virtually identical for both age groups. Indeed, no statistically significant differences between age groups were found in the cumulative percentage of words correctly identified at any band value (for all independent samples t-tests, $p>0.4$).

For each list, we calculated each individual's threshold (the number of bands at which the cumulative percentage of correctly identified words was 50%). For both age groups, word identification improved similarly across the four lists, with mean band thresholds of 5.8, 5.3, 5.0, and 4.9 (overall mean threshold=5.25 bands) for the younger group, and 5.5, 5.3, 5.1, and 5.0 (overall mean threshold=5.25 bands) for the older group. This description was confirmed by an analysis of variance (ANOVA) with age as a between-subjects factor and list order as a within-subjects factor that revealed no significant effect of age, $F(1,22)=0.000$, $p>0.9$, but a significant effect of list order, $F(3,66)=14.53$, $p<0.001$, with no significant interaction between age and list order, $F(3,66)=1.65$, $p>0.1$. A Tukey test of multiple comparisons confirmed that performance did not differ significantly between Lists 1 and 2 ($p>0.1$), but that there was significant improvement between List 1 and List 3 ($p<0.01$) and between List 1 and List 4 ($p<0.002$), although performance on Lists 3 and 4 was not significantly different ($p>0.5$). Thus, there is an overall improvement in performance with increasing exposure to noise-vocoded stimuli across lists, but there does not seem to be an age-related difference in word identification or an age-related difference in the degree of improvement across lists.

III. EXPERIMENT 2

The goal of Experiment 2 was to investigate if the older and younger participants who had gained experience with noise-vocoded speech in Experiment 1 would differ in their ability to identify noise-vocoded words using a more common method of presentation in which the number of bands used in vocoding was blocked rather than gated as it had been in Experiment 1.

A. Method

1. Participants

The participants were those who completed Experiment 1.

2. Stimuli and apparatus

The test stimuli were the digital recordings of the four W-22 word lists that have been standardized for use in clinical speech audiometry (see Penrod, 1985) and distributed on compact disk by Auditec of St. Louis. Like the NU-6 lists, each W-22 list consists of 50 monosyllabic words spoken by a male talker preceded with the carrier phrase "*Say the word*…." The NU-6 lists were designed after the W-22 lists to be a more sensitive test for individuals with predominantly high-frequency hearing loss. The main differences between the NU-6 and W-22 lists are that the word frequency profile is lower and the occurrence of high-frequency consonants is higher in the NU-6 lists than in the W-22 lists. The digitized recordings were subjected to the same noise-vocoding procedure used by Eisenberg *et al.* (2000) and described earlier. For each list, the words were noise-vocoded using 16, 8, 4, and 2 bands. Stimuli were delivered using the same apparatus as was used in Experiment 1.

3. Procedure

Immediately following Experiment 1, in which the participants were familiarized with noise-vocoded speech, the participants completed Experiment 2. In contrast to Experiment 1, the presentation of stimulus conditions was blocked in Experiment 2, with the order of conditions progressing from easiest to hardest. First, participants heard a word list in the 16-band condition. Next, they heard a word list in the 8-band condition, then a word list in the 4-band condition,

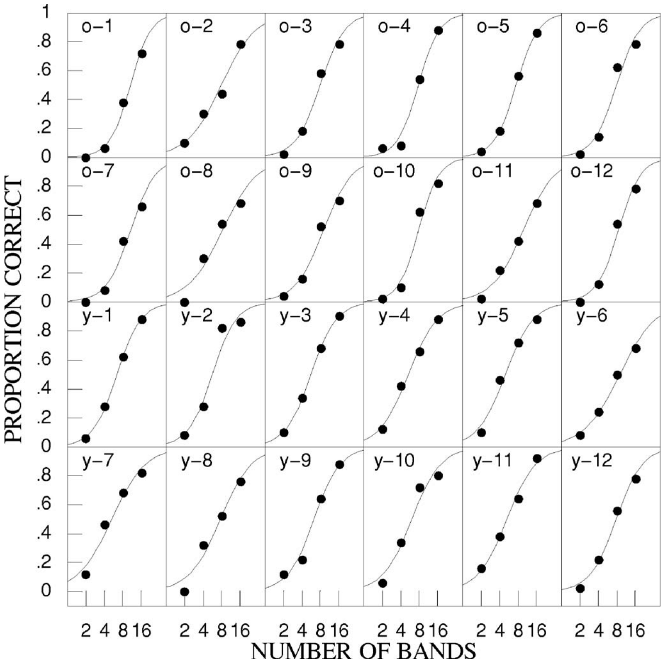

FIG. 3. The proportion of words correctly identified as a function of the number of bands for individual younger (y) and older (o) participants in Experiment 2.

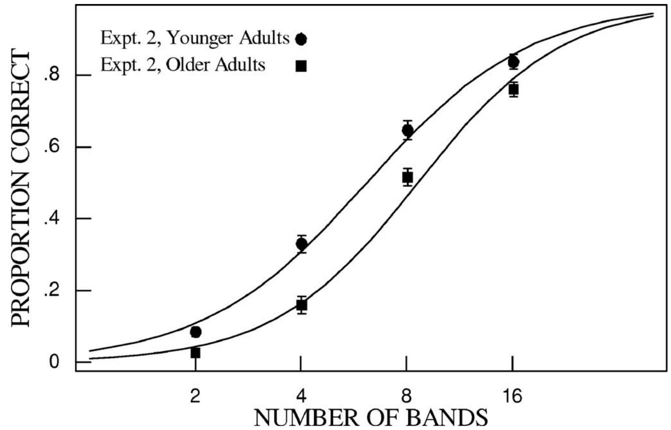

FIG. 4. The proportion of words correctly identified, averaged across participants, as a function of the number of bands for younger (circles) and older (squares) participants in Experiment 2. Standard error bars are shown.

and finally a word list in the 2-band condition. Words within a list were presented in random order and the list order was counterbalanced across participants so that each list was presented an equal number of times in each order position for each group.

For each word, the participant was asked to identify the word or respond "I don't know." Guessing was strongly encouraged. The experimenter, who was outside the booth, listened to and immediately scored each response. If the experimenter was uncertain about any response, the participant was asked to repeat and spell the word aloud. Any difference between response and target was marked as an error. No feedback was given after a response. All sessions were audiotaped. The testing session lasted approximately 25 min.

B. Results and discussion

Figure 3 plots the percentage of words that were correctly identified as a function of the number of bands for each participant. Logistic functions of the form

$$y = \frac{1}{1 + e^{-\sigma[(\log_{10} x) - \mu]}}$$

were fit to the data of each participant, where y is the proportion of words correctly identified, x is the number of bands, μ is the threshold (the value of log x resulting in 50% correct identification of the words), and σ is the slope parameter of the psychometric function (see the Appendix). It is apparent in Fig. 3 that a logistic function provides a good fit to the data of each individual.

Figure 4 plots the average proportion of words correctly identified by the younger and older groups by band condition. Figure 4 suggests that the psychometric functions for younger and older adults have similar slopes (σ), but different threshold values (μ). That is, the psychometric function for younger adults is the same as that for older adults, except that it is shifted to the left by 0.14 log units. To verify this, we conducted two ANOVAs with pure-tone thresholds (at all eight audiometric test frequencies) as covariates: one for the values of slopes and another for the band thresholds obtained from the individual data of younger and older adults (Fig. 3). These tests indicate that younger and older adults differed with respect to band thresholds, with a significant main effect of age, $F(1,14)=5.104$, $p<0.05$, but not with respect to slopes, $F(1,14)=0.814$, $p>0.1$. As shown in Fig. 4, older adults needed an average of 8.55 bands to correctly identify 50% of the noise-vocoded words, whereas younger adults needed an average of only 6.13 bands to achieve the same level of performance.

It is interesting that the band thresholds calculated in Experiment 2 based on 50% correct word identification were larger for both groups (older=8.55; younger=6.13), compared to the band thresholds calculated in Experiment 1 based on 50% cumulative correct word identification (5.25 for both age groups). Experiment 1 should have been more challenging than Experiment 2 insofar as the participants were less experienced with noise-vocoded speech in Experiment 1 than in Experiment 2, and the NU-6 word lists used in Experiment 1 could have been more difficult than the W-22 word lists used in Experiment 2, especially for the older listeners with high-frequency hearing loss because of the emphasis in the NU-6 lists on high-frequency consonants. The differences in the size of the band thresholds in the two experiments seem more likely to be explained by the way in which the functions were measured than by the stimuli that were used.

More important, the finding of a significant effect of age on the number of bands needed to correctly identify 50% of the noise-vocoded words differs from the finding of no effect of age on performance in Experiment 1. It is possible that the discrepancy between the finding of no age-related difference in Experiment 1 but a significant age-related difference in Experiment 2 could be explained by methodological differences between the two experiments, including differences associated with varying the number of bands using gating versus blocked presentation methods, or differences related to whether or not feedback was provided. Another possibility is that there were age-related differences in the carry-over of learning from Experiment 1 to Experiment 2. Age-related differences in carry-over of learning would be consistent with the findings of previous studies showing that perceptual learning gained from a training session was more beneficial to younger than to older adults in a subsequent testing session (e.g., Peelle and Wingfield, 2005; Sommers, 1997). According to this explanation, the reason that an age effect was found in Experiment 2, but not in Experiment 1, would be that the younger adults were better able than the older adults to generalize or to use the experience they had gained during Experiment 1 to aid their performance in Experiment 2. In other words, there might be no age-related difference in temporal processing of envelope cues per se, but only a difference in the ability of younger and older adults to generalize from the training set to a testing set.

IV. EXPERIMENT 3

In Experiment 3, we tested new groups of younger and older listeners using the same procedure as in Experiment 2; however, the participants in Experiment 3 had never heard noise-vocoded stimuli prior to the experiment. To evaluate the possibility that the age-related differences found in Experiment 2 were due to age-related differences in the carry-over of learning from Experiment 1, the performance of the younger and older listeners in Experiment 2 (who had experience with noise-vocoded stimuli in Experiment 1) was compared to that of the younger and older listeners in Experiment 3 (who had not been previously exposed to noise vocoding).

A. Method

1. Participants

Twelve younger adults (mean age=21.0, s.d.=2.9, range=17–25 years) and twelve older adults (mean age =67.4, s.d.=2.8, range=64–72 years) participated in this experiment. Both groups had pure-tone air-conduction thresholds in the test ear that were less than or equal to 25 dB HL from 0.25 to 3 kHz (Table I). All participants had learned English before the age of 5 years and had been educated in English in a country where it is the dominant language. Mean scores on the Mill-Hill Vocabulary Scale (Raven, 1965) were better for older than for younger participants [mean for the younger group=12.7, s.d.=2.3; mean for the older group=15.0, s.d.=1.7; $t(22)=-2.85$, $p<0.05$]. On average, the younger adults had 1.2 years more education than the older adults [mean for younger participants=15.3 years of education, s.d.=3.1; mean for older participants=14.1, s.d.=2.2 years); $t(22)=1.06$, $p>0.10$]. All of the participants were paid volunteers recruited from the local community who gave informed consent in compliance with the protocol approved by the university's ethics review board. No participant had previously listened to noise-vocoded speech.

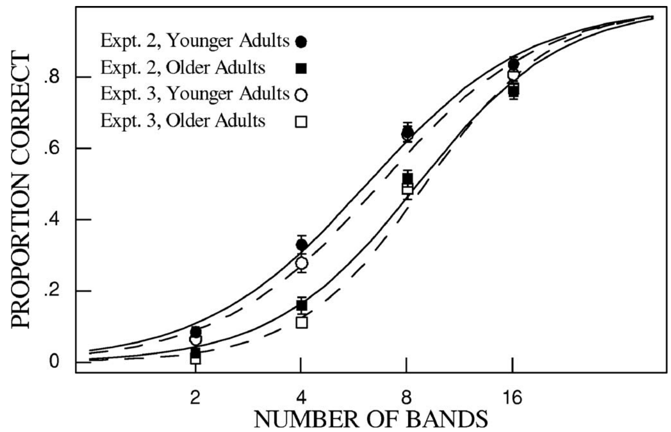

FIG. 5. The proportion of words correctly identified, averaged across participants, as a function of the number of bands for younger (circles) and older (squares). Closed symbols and solid lines are for Experiment 2. Open symbols and dashed lines are for Experiment 3. Standard error bars are shown.

2. Stimuli and procedure

The same materials and procedure used in Experiment 2 were used in Experiment 3.

B. Results and discussion

Figure 5 plots the average proportion of words correctly identified by the younger and older participants in Experiment 3 compared to the data obtained in Experiment 2. As in Experiment 2, the psychometric functions for younger and older adults have similar slopes (σ) and threshold values (μ). To verify this pattern for Experiment 3, we conducted two ANOVAs with pure-tone thresholds (at all eight audiometric test frequencies) as covariates: one for the values of slopes and another for the band thresholds obtained from the individual data of younger and older adults. These tests indicate that younger and older adults differed with respect to band thresholds, with a significant main effect of age, $F(1,14)=9.074$, $p<0.01$, but not with respect to slopes, $F(1,14)=1.295$, $p>0.1$. Older adults needed an average of 8.97 bands to correctly identify 50% of the noise-vocoded words, whereas younger adults needed an average of only 6.67 bands to achieve the same level of performance.

Furthermore, the similarity of the functions in Fig. 5 suggests that previous experience with noise-vocoded words had little effect on the threshold values (μ) for either age group. We conducted an ANOVA on the threshold and slope values with experiment (Experiment 2 versus Experiment 3) and age (younger versus older) as between-subjects factors. For the band threshold values, there was a significant difference between younger and older adults, $F(1,44)=47.47$, $p<0.0001$, but no significant difference between experiments, $F(1,44)=2.07$, $p>0.05$, and no significant interaction between age and experiment, $F(1,44)=0.15$, $p>0.05$. For the slope values, there was a significant difference between age groups, $F(1,44)=20.53$, $p<0.001$, and a borderline significant difference between experiments, $F(1,44)=4.32$, $p<0.05$, but no significant interaction between age and experiment, $F(1,44)=2.07$, $p>0.05$. Crucially, prior exposure and learning did not significantly alter the band threshold values for either the younger or the older adults and the effect of age on word identification performance was observed in both Experiments 2 and 3. Therefore, we cannot attribute the results found in Experiment 2 to an age-related decline in the ability to carry-over learning gained in Experiment 1.

V. GENERAL DISCUSSION

The primary goal of the present study was to examine the effect of age on the ability of listeners to use envelope cues to identify words spoken in a carrier phrase when fine structure cues are minimized in noise-vocoded speech. A secondary goal was to investigate how age may affect perceptual learning as listeners acquire and consolidate learning gained from experience listening to noise-vocoded speech. In Experiment 1, we found no significant difference between the word identification accuracy of younger and older listeners when noise-vocoded stimuli were presented using an adapted gating procedure in which each target word was presented with an increasing number of bands until it was correctly identified. Not only was the mean band threshold for 50% cumulative correct word identification identical for both age groups, but the entire functions of the cumulative percentage of correctly identified words by band condition were also nearly identical for the two age groups. Both age groups also demonstrated similar perceptual learning in Experiment 1. In contrast, when the same listeners were subsequently tested in Experiment 2, there was a significant age-related difference in band threshold when each target word was presented in a blocked design. Similar to the results of the participants in Experiment 2 who had been exposed to noise-vocoded speech in Experiment 1, for the participants in Experiment 3 who had no prior exposure to noise-vocoded speech, older adults performed significantly worse than younger adults. Thus, age-related differences in the carry-over of learning from Experiment 1 do not seem to account for the discrepancy between the finding of an age effect in Experiment 2 but not in Experiment 1.

In this discussion, we will compare our results to those reported in previous research. Then we will consider if the discrepancy between Experiments 1 and 2 could be explained by methodological differences between the two experiments, namely differences associated with varying the number of bands when using gating versus blocked presentation methods or differences related to whether or not feedback was provided. Finally, we will discuss the absence of the carry-over of learning from Experiment 1 to Experiment 2.

A. Identification of noise-vocoded speech

To our knowledge, no other studies have adapted the gating paradigm to noise vocoding as we did in Experiment 1. Previous noise-vocoding studies typically used a blocked design similar to the one we used in Experiment 2 (e.g., Shannon *et al.*, 1995; Dorman *et al.*, 1998; Loziou *et al.*, 1999; Eisenberg *et al.*, 2000; Fu and Nogaki, 2005; Souza and Boike, 2006), or they used stimuli in only one selected band-processed condition (e.g., Davis *et al.*, 2005; Trout, 2005). In previous noise-vocoding studies, the accuracy with which the vocoded words were identified by younger adults has often been higher than was found in our Experiments 2 and 3. This difference may best be explained by differences in training and the type of speech material. We provided either minimal training (Experiment 2) or no training (Experiment 3) with noise-vocoded stimuli, whereas others have provided extensive training. For example, to achieve near-perfect levels of word identification with only four bands, Shannon and colleagues (1995) provided 8–10 h of training. Furthermore, the open-set test of monosyllabic words we used was more demanding than the closed-set recognition tests used in many previous studies (e.g., van Tasell *et al.*, 1992; Souza and Boike, 2006). In addition, other studies likely achieved higher levels of performance in low band-processed conditions because they used sentences in which lexical knowledge could be used to help decipher the degraded signal (e.g., Fishman *et al.*, 1997), whereas the carrier phrase we used did not provide listeners with any opportunity to use context to advantage. Results more similar to ours have been reported in studies in which the stimuli and response alternatives were also more similar to those of our study. In studies that tested monosyllabic word identification using an open set, results have ranged from about 55% correct in a 4-band processed condition (Friesen *et al.*, 2001) down to only 9.9% accuracy for 5-band noise-vocoded words that varied in lexical difficulty and were presented with no carrier phrase (Trout, 2005). Importantly, although identification of noise-vocoded words may have been harder in the present experiments than in less challenging tests used in prior studies, we were able to detect both similarities and differences between younger and older listeners.

One prior study has examined the relative effects of age and degree of hearing loss on the ability of adults to process noise-vocoded speech (Souza and Boike, 2006); however, their sample did not include older adults with good audiograms. Nevertheless, our finding of an age effect in Experiment 2 is in line with the results of Souza and Boike (2006) who found that age, but not degree of hearing loss, was a significant predictor of the ability of listeners to identify noise-vocoded /aCa/ nonsense bisyllables in a 16-alternative closed-choice task. They interpreted their findings for adults ranging in age from 23 to 80 years and degree of hearing loss from mild to severe as evidence for an age-related deficit in the use of temporal envelope information across all band conditions (1-, 2-, 4-, and 8-band conditions). Both age groups in our present study had clinically normal audiograms in the speech range, although the mean thresholds of the older adults were higher than those of the younger adults, especially at the highest frequencies (4–8 kHz). Nevertheless, consistent with the conclusion of Souza and Boike (2006) that there is an age-related deficit that is not explained by degree of hearing loss, the smaller audiometric threshold differences between the younger and older adults in our study could not explain why we found significant group differences in word identification performance in Experiments 2 and 3. It also seems unlikely that audiometric threshold differences would have affected performance in Experiments 2 and 3, but not in Experiment 1. If high-frequency audiometric loss at frequencies of 4 kHz and higher contributed to the problems of the older adults then age-related differences should have been more pronounced in Experiment 1 than in Experiments 2 and 3 because the NU-6 word lists used in Experiment 1 were designed to be more challenging for people with high-frequency hearing loss. In fact, using the proportion of words correctly identified in Experiment 1 as a baseline measure of word identification, both age groups achieved near-ceiling performance and the mean proportion correct for younger adults (0.97, s.d.=0.023) and older adults (0.96, s.d.=0.042) did not differ significantly, $t(22)=1.25$, $p>0.10$. It is also worth noting that the closeness of the cumulative percent correct functions for the two age groups shown in Fig. 2 seems inconsistent with the possibility that age-related differences were found only in Experiment 2 because older adults were less willing than younger adults to guess the first time that a word was presented. Non-audiometric age-related differences may better explain the pattern of results.

B. Age-related differences in temporal processing

The main finding from Experiments 2 and 3 is that there is an age-related reduction in ability to use envelope cues to identify noise-vocoded words, as illustrated by the similar slopes (σ), but significantly different band thresholds (μ) of the younger and older adults. Although much prior research has been interpreted as suggesting that the processing of prosody is largely preserved with age, a closer examination of the pattern of results of key studies demonstrates that although older adults are able to benefit from prosodic cuing, including envelope duration and shape cues, it is not the case that older adults achieve the same level of performance as is achieved by younger adults. For example, in one study of speech prosody, the standard time-gating technique was used to investigate age-related differences in the use of envelope cues for word identification in three experimental conditions: in one condition, only the onset of the word was presented; in another condition, the onset of the word was presented and noise was used to terminate the word, adding information about the duration of the word; in the remaining condition, the onset of the word was provided plus a noise shaped by the speech envelope that provided duration and additional prosodic information (Wingfield *et al.*, 2000). Although older and younger listeners benefited similarly from the addition of duration and envelope cues, the younger adults outperformed the older adults even in the onset plus envelope-shaped noise condition. Thus, the age-related difference in overall performance that was observed seems to be consistent with our finding of an age-related difference in word identification based on the use of envelope cues in noise-vocoded speech.

Previous research has found evidence of age-related temporal processing deficits related to subsegmental and segmental speech cues. Our results suggest that there are also age-related differences in the use of envelope cues that could be relevant for processing suprasegmental speech information. Whether or how the age-related differences in auditory temporal processing relevant to the different levels of speech information are related has yet to be determined. Importantly, the present study establishes the possibility that age-related differences in auditory temporal processing exist at the suprasegmental level even when audiometric thresholds in the speech range are within clinically normal limits.

C. Benefit from repetition and feedback

Contrary to the findings from Experiments 2 and 3, in Experiment 1 we found no age-related differences in ability to use envelope cues when identifying noise-vocoded words. Nonetheless, the absence of an effect of age on word identification in Experiment 1 seems to be consistent with the idea that the processing of prosodic cues is relatively preserved in older adults. Clearly, it is important to consider why the results for Experiments 1 and 2 are discrepant. A number of methodological differences, including differences in the stimuli, may provide an explanation for the discrepancies. We will consider repetition and feedback as two such possibilities.

When words are presented using the band-gating procedure, as they were in Experiment 1, there is an opportunity for the listener to benefit from the summing of information that is incremented as the number of bands is increased over sequential presentations. No benefit from this type of summed information is available when a stimulus is presented only a single time, as was done in Experiment 2. The use of summed information has been cited in psychophysical studies to explain improvements in performance on tasks such as detecting interaural differences in intensity for trains of clicks (Hafter and Dye, 1983; Hafter *et al.*, 1983). Furthermore, it seems possible that summation of information over repeated utterances could be at play in everyday conversational behavior. Repetition, especially repetition with clearer pronunciation, is the most common conversational repair strategy (Drew, 1997). For a listener in a conversation, the information about the identity of a word provided by its first presentation may be used in conjunction with the information gained from later repetition to achieve correct word identification. That is, it could be possible to use the first degraded exposure to the utterance to narrow the set of possible lexical alternatives and to focus listening on the critical but missed portions of the utterance when it is repeated. Future experiments could directly explore the possible contribution of stimulus repetition.

If listeners can sum information to aid word identification, since everyday listening environments are more challenging for older adults than for younger adults, older adults are likely to have considerable experience in using this type of compensatory mechanism. Age-related differences may be found in Experiment 2 because there is no opportunity to compensate by summing information. The notion that older adults may sum information to support word identification in adverse listening conditions is consistent with other research indicating that older adults are better able to use other compensatory mechanisms, such as phonological knowledge (e.g., Pichora-Fuller *et al.*, 2006) or sentence context to identify speech in adverse listening situations (e.g., Pichora-Fuller *et al.*, 1995; Wingfield *et al.*, 2005). More generally, these findings are consistent with the idea that age-related differences on a range of cognitive measures are reduced when aspects of the environment can be used to support performance on a demanding task (e.g., Craik, 1983, 1986). Convergent evidence of age-related compensation during perceptual and cognitive tasks has also been found in cognitive neuroscience research showing that there is more bilateral activation of the brain for older adults when they achieve the same performance as younger adults on a variety of tasks (e.g., Grady, 2000; Cabeza, 2002; Reuter-Lorenz, 2002).

An alternative explanation for the differences related to the presentation method used in the two experiments is that there was feedback provided in Experiment 1, but not in Experiment 2. In particular, the feedback provided in Experiment 1 may account for the improvement from List 1 to List 4 that was observed for both age groups, in agreement with previous studies demonstrating younger adults' abilities to learn to identify noise-vocoded speech (Davis *et al.*, 2005). Without the benefit from feedback in Experiment 2, our older participants may have been more disadvantaged than their younger counterparts. The idea that older adults may be facilitated differentially by feedback has been used to explain findings in a study of the effect of age on auditory lexical decision (Stine-Morrow *et al.*, 1999).

D. Carry-over of learning

Neither younger adults nor older adults seemed to be able to carry-over learning from Experiment 1 to improve their performance on Experiment 2, as seen by the lack of a significant difference between the mean band thresholds of the experienced (Experiment 2) and inexperienced groups (Experiment 3). An obvious explanation for this null effect is that there was not enough training. Perceptual training studies typically involve training sessions that take many hours over the course of many days whereas the training session in the current study lasted, at most, 2 h (Kraus *et al.*, 1995; Tremblay *et al.*, 2001). Nevertheless, in the current study, there was perceptual learning or familiarization within the training set for both age groups, as evidenced by the significant improvement over list presentations in Experiment 1. Furthermore, both age groups seemed to reach their plateau performance insofar as there was no significant difference for either group between their performance on Lists 3 and 4. The improvement of the participants in Experiment 1 is consistent with studies citing short-term perceptual adaptation as the explanation for improved performance in identification of distorted speech stimuli (Clarke, 2002; Mehler *et al.*, 1993; Davis *et al.*, 2005; Peelle and Wingfield, 2005).

Another possible explanation for the lack of an effect of learning is that learning does not carry-over across the different presentation methods used in Experiments 1 and 2. In the present study, the gating and feedback used in Experiment 1 differed in presentation from the single, blocked presentation of stimuli in Experiment 2. Although both experiments used simple, monosyllabic words, the words were not identical. In addition, the extent of experience with the bands tested in Experiment 2 was not uniform in Experiment 1 because of the differences in the gating and blocked organization of the presentation of the words. One or more of these differences between the experiments could have prevented carry-over. The idea that perceptual learning is task-specific and that training effects may not generalize has been suggested in other auditory learning research. For example, Irvine *et al.* (2000) found that the effect of training did not transfer to other frequencies for a frequency discrimination task. Similarly, Burk *et al.* (2006) found that training on isolated words presented in noise was not sufficient to yield large improvements on untrained words presented in noise. In the visual domain, Fahle and Morgan (1996) found no transfer of perceptual learning or training between similar stimuli tested in two different tasks. Our finding of no difference between the word identification thresholds of experienced listeners in Experiment 2 and inexperienced listeners in Experiment 3 may be in keeping with the more general finding that perceptual learning can be highly task-specific and does necessarily generalize to other tasks given only prior exposure to similar stimuli in a different context/task (Cohen *et al.*, 2006).

In summary, we observed perceptual learning in Experiment 1, but no significant carry-over to Experiment 2 for either younger or older adults. Many differences between the experiments may have prevented carry-over of learning. In any case, the lack of carry-over of learning does not seem to explain why there were age-related differences in Experiment 2 since the same pattern of results was found for an inexperienced group who completed the same tests in Experiment 3.

VI. CONCLUSIONS

Without the benefit of summing information from repetitions of a word and/or benefit from feedback, older adults do not use envelope cues as well as younger adults to identify noise-vocoded words in an open-set task. This new evidence that older adults show a deficit in using the temporal amplitude envelope cues relevant to suprasegmental aspects of speech perception extends previous evidence that there are age-related declines in other aspects of auditory temporal processing relevant to other levels of speech processing. Importantly, we also found that when noise-vocoded speech is presented with the opportunity to sum information in the speech signal over repetitions and/or feedback is provided after each presentation, as in Experiment 1, then age-related differences are eliminated. Lastly, we found that perceptual learning with noise-vocoded stimuli occurred similarly for both age groups, but did not generalize across experiments for either age group, suggesting that age-related differences in carry-over of learning did not explain the age-related differences that we observed in ability to use envelope cues to identify noise-vocoded words.

ACKNOWLEDGMENTS

The authors are grateful to Ewen MacDonald for preparing the noise-vocoding program. This research was funded by the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada.

APPENDIX

We can test whether or not the logistic function provides a good fit to the data using the normalized Pearson's χ^2 statistic where

$$\text{norm } \chi^2 = \sum_{i=1}^{n} \frac{\left(y_i - \frac{1}{1+e^{-\sigma[(\log_{10} x_i)-\mu]}}\right)^2}{\frac{1}{1+e^{-\sigma[(\log_{10} x_i)-\mu]}}} + \sum_{i=1}^{n} \frac{\left(\frac{1}{1+e^{-\sigma[(\log_{10} x_i)-\mu]}} - y_i\right)^2}{1-\frac{1}{1+e^{-\sigma[(\log_{10} x_i)-\mu]}}}.$$

x_i is the number of bands, y_i is the proportion of correctly identified words at x_i, μ, and σ are the threshold and slope parameters, respectively, of the logistic function, and n_i is the number of stimuli in the experiment. Specifically, the values of μ and σ are systematically varied to minimize the normalized χ^2 statistic. When the normalized χ^2 statistic is multiplied by the number of times, N, each stimulus is presented in the experiment, it becomes a Pearson's χ^2 statistic with $n-2$ degrees of freedom. In Experiment 2, $N=50$, and $n=4$. Hence the degrees of freedom are $4-2=2$.

Because the test statistic is distributed according to χ^2 with 2 degrees of freedom, we can test the null hypothesis to determine whether or not the logistic function provides a good fit to the data from each individual. All tests were conducted using a Bonferroni correction for the number of tests conducted (in this case 24 tests were conducted to see if the logistic function could describe individual data). Hence, to correct for the number of tests we used $\alpha=0.05/24=0.0021$ for each test. Using the Bonferroni correction, we failed to reject the null hypothesis for all of the participants, and therefore concluded that the logistic function provided a good fit to the individual data.

Abel, S. M., Krever, E. M., and Alberti, P. W. (**1990**). "Auditory detection, discrimination and speech processing in aging, noise-sensitive and hearing impaired listeners," Scand. Audiol. **19**, 43–54.

Alain, C., McDonald, K. L., Ostroff, J. M., and Schneider, B. (**2001**). "Age-related changes in detecting a mistuned harmonic," J. Acoust. Soc. Am. **109**, 2211–2216.

Bergerson, T. R., Schneider, B. A., and Hamstra, S. J. (**2001**). "Duration discrimination in younger and older adults," Can. Acoust. **29**, 3–9.

Burk, M. H., Humes, L. E., Amos, N. E., and Strauser, L. E. (**2006**). "Effect of training in word-recognition performance in noise for young normal-hearing and older hearing-impaired listeners," Ear Hear. **27**, 263–278.

Cabeza, R. (**2002**). "Hemispheric asymmetry reduction in older adults: The HAROLD model," Psychol. Aging **17**, 85–100.

Callan, D. E., Taijma, K., Callan, A. M., Kubo, R., Masaki, S., and Akahane-Yamada, R. (**2003**). "Learning-induced neural plasticity associated with improved identification performance after training difficult second-language phonetic contrasts," Neuroimage **19**, 113–124.

CHABA (Committee on Hearing, Bioacoustics, and Biomechanics) Working Group on Speech Understanding and Aging, National Research Council (**1988**). "Speech Understanding and Aging," J. Acoust. Soc. Am. **83**, 850–805.

Clarke, C. M. (**2002**). "Perceptual adjustment to foreign-accented English with short-term exposure," Proceedings of the Seventh International Conference on Spoken Language Processing, Denver, Co., Vol. **1**, pp. 253–256.

Cohen, Y. E., Russ, B. E., and Lee, Y. S. (**2006**). "Neural and behavioural correlates of auditory categorization," Programme and Abstracts for the International Conference on the Auditory Cortex, MRC Institute of Hearing Research, Nottingham, UK, pp. 53–54.

Craik, F. I. M. (**1983**). "On the transfer of information from temporary to permanent memory," Philos. Trans. R. Soc. London, Ser. B **302**, 341–359.

Craik, F. I. M. (**1986**). "A functional account of age differences in memory," in *Human Memory and Cognitive Capabilities, Mechanisms, and Performances*, edited by F. Klix and H. Hagendorf (Elsevier Science Publishers, North-Holland, Amsterdam), pp. 499–522.

Culter, A., Dahan, S., and van Donselaar, W. (**1997**). "Prosody in the comprehension of spoken language: A literature review," Lang. Speech **40**, 141–201.

de Boer, E., and Dreschler, W. A. (**1987**). "Auditory psychophysics: Spectrotemporal representation of signals," **38**, 181–202.

de Cheveigné, A. (**2003**). "Time domain auditory processing of speech," J. Phonetics **31**, 547–461.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., and McGettigan, C. (**2005**). "Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences," J. Exp. Psychol. **134**, 222–241.

Divenyi, P., and Simon, H. (**1999**). "Hearing in aging: Issues old and young," Current Opinion in Otolaryngology and Head & Neck Surgery **7**, 282–289.

Dorman, M. F., Loizou, P. C., Fitzke, J., and Tu, Z. (**1998**). "The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels," J. Acoust. Soc. Am. **104**, 3583–3585.

Drew, P. (**1997**). " 'Open' class repair initiators in response to sequential sources of troubles in conversation," J. Pragmatics **28**, 69–101.

Dupoux, E., and Green, K. (**1997**). "Perceptual adjustment to highly compressed speech: Effects of talker and rate changes," J. Exp. Psychol. **33**, 914–927.

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., and Boothroyd, A. (**2000**). "Speech recognition with reduced spectral cues as a function of age," J. Acoust. Soc. Am. **107**, 2704–2710.

Fahle, M., and Morgan, M. (**1996**). "No transfer of perceptual learning between similar stimuli in the same retinal position," Curr. Biol. **6**, 292–297.

Fishman, K., Shannon, R. V., and Slattery, W. H. (**1997**). "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," J. Speech Hear. Res. **40**, 1201–1215.

Fitzgibbons, P., and Gordon-Salant, S. (**1994**). "Age effects on measures of auditory duration discrimination," J. Speech Hear. Res. **37**, 662–670.

Fitzgibbons, P. J., and Gordon-Salant, S. (**1995**). "Age effects on duration discrimination with simple and complex stimuli," J. Acoust. Soc. Am. **98**, 3140–3145.

Fitzgibbons, P. J., and Gordon-Salant, S. (**1996**). "Auditory temporal processing in elderly listeners," J. Am. Acad. Audiol **7**, 183–189.

Friesen, L., Shannon, R., Baskent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Am. **110**, 1150–1163.

Frisina, R. D. (**2001**). "Possible neurochemical and neuroanatomical bases of age-related hearing loss-presbycusis," Semin. Hear. **22**, 213–225.

Fu, Q. J., and Nogaki, G. (**2005**). "Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing," **6**, 19–27.

Gordon-Salant, S., and Fitzgibbons, P. J. (**1993**). "Temporal factors and speech recognition performance in young and elderly listeners," J. Speech Hear. Res. **36**, 1276–1285.

Gordon-Salant, S., and Fitzgibbons, P. J. (**1997**). "Selected cognitive factors and speech recognition performance among young and elderly listeners," J. Speech Hear. Res. **40**, 423–431.

Gordon-Salant, S., Yeni-Komishian, G. H., Fitzgibbons, P. J., and Barrett, J. (**2006**). "Age-related differences in identification and discrimination of temporal cues in speech segments," J. Acoust. Soc. Am. **119**, 2455–2466.

Grady, C. L. (**2000**). "Functional brain imaging and age-related changes in cognition," Biol. Psychol. **54**, 259–281.

Greenberg, S. (**1996**). "Auditory processing of speech," in *Principles of Experimental Phonetics*, edited by N. J. Lass (Mobsy, St. Louis), pp. 362–407.

Greenspan, S. L., Nusbaum, H. C., and Pisoni, D. B. (**1988**). "Perceptual learning of synthetic speech produced by rule," J. Exp. Psychol. Learn. Mem. Cogn. **17**, 152–162.

Greenwood, D. D. (**1990**). "A cochlear frequency-position function for several species—29 years later," J. Acoust. Soc. Am. **87**, 2592–2605.

Grosjean, F. (**1996**). "Gating," Lang. Cognit. Processes **11**, 597–604.
Hafter, E. R., and Dye, R. H., Jr. (**1983**). "Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number," J. Acoust. Soc. Am. **73**, 644–651.
Hafter, E. R., Dye, R. H., Jr., and Wenzel, E. (**1983**). "Detection of interaural differences of intensity in trains of high-frequency clicks as a function of interclick interval and number," J. Acoust. Soc. Am. **73**, 1708–1713.
Irvine, D. R. F., Martin, R. L., Klimkeit, E., and Smith, R. (**2000**). "Specificity of perceptual learning in a frequency discrimination task," J. Acoust. Soc. Am. **108**, 2964–2968.
Konkle, D. F., Beasley, D. S., and Bess, F. H. (**1977**). "Intelligibility of time-altered speech in relation to chronological aging," J. Speech Hear. Res. **20**, 108–115.
Kraus, N., McGee, T., Carrell, T., King, C., and Tremblay, K. (**1995**). "Central auditory system plasticity associated with speech discrimination training," J. Cogn Neurosci. **7**, 27–34.
Loizou, P. C., Dorman, M., and Tu, Z. (**1999**). "On the number of channels needed to understand speech," J. Acoust. Soc. Am. **106**, 2097–2103.
Mehler, J., Sebastian, N., Altmann, G., Dupoux, E., Christophe, A., and Pallier, C. (**1993**). "Understanding compressed sentences—The role of rhythm and meaning," Ann. N.Y. Acad. Sci. **682**, 272–282.
Peelle, J. E., and Wingfield, A. (**2005**). "Dissociable components of perceptual learning revealed by adult age differences in adaptation to time-compressed speech," J. Exp. Psychol. Hum. Percept. Perform. **31**, 1315–1330.
Penrod, J. P. (**1985**). "Speech discrimination testing," in *Handbook of Clinical Audiology*, 4th ed., edited by J. Katz (Williams and Wilkins, Baltimore, MD), pp. 235–255.
Phillips, D. P. (**1995**). "Central auditory processing: A view from auditory neuroscience," Am. J. Otol. **16**, 338–352.
Pichora-Fuller, M. K., and Schneider, B. A. (**1992**). "The effect of interaural delay of the masker on masking-level difference in young and old adults," J. Acoust. Soc. Am. **91**, 2129–2135.
Pichora-Fuller, M. K., Schneider, B., Benson, N., Hamstra, S. J., and Storzer, E. (**2006**). "Effect of age on gap detection in speech and non-speech stimuli varying in marker duration and spectral symmetry," J. Acoust. Soc. Am. **119**, 1143–1155.
Pichora-Fuller, M. K., Schneider, B., MacDonald, E., Brown, S., and Pass, H. (**2007**). "Temporal jitter disrupts speech intelligibility: A simulation of auditory aging," Hear. Res. **223**, 114–121.
Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (**1995**). "How young and old adults listen to and remember speech in noise," J. Acoust. Soc. Am. **97**, 593–608.
Pichora-Fuller, M. K., and Souza, P. (**2003**). "Effects of aging on auditory processing of speech," Int. J. Audiol. **42**, S11–S16.
Raven, J. C. (**1965**). *The Mill Hill Vocabulary Scale* (Lewis, London).
Reuter-Lorenz, P. A. (**2002**). "New visions of the aging mind and brain," Trends Cogn. Sci. **6**, 394–400.
Rosen, S. (**1992**). "Temporal information in speech: acoustic, auditory and linguistic aspects," Philos. Trans. R. Soc. London, Ser. B **336**, 367–373.
Schneider, B. A., Daneman, M., and Murphy, D. R. (**2005**). "Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing?," Psychol. Aging **20**, 261–271.
Schneider, B. A., and Pichora-Fuller, M. K. (**2001**). "Age-related changes in temporal processing: Implications for listening comprehension," Semin. Hear. **22**, 227–239.
Schneider, B. A., Pichora-Fuller, M. K., Kowalochuk, D., and Lamb, M. (**1994**). "Gap detection and the precedence effect in younger and older adults," J. Acoust. Soc. Am. **95**, 980–991.
Shannon, R. V. (**2002**). "The relative importance of amplitude, temporal and spectral cues for cochlear implant processor design," Amer. J. Audiology **11**, 124–127.
Shannon, R. V., Zeng, F., Kamath, V., and Wygonski, J. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.
Snell, K. B. (**1996**). "Age-related changes in temporal gap detection," J. Acoust. Soc. Am. **101**, 2214–2220.
Snell, K. B., and Frisina, R. D. (**2000**). "Relationships among age-related differences in gap detection and word recognition," J. Acoust. Soc. Am. **107**, 1615–1626.
Sommers, M. S. (**1997**). "Stimulus variability and spoken word recognition. II. The effects of age and hearing impairment," J. Acoust. Soc. Am. **101**, 2278–2288.
Souza, P. E., and Boike, K. T. (**2006**). "Combining temporal-envelope cues across channels: Effects of age and hearing loss," J. Speech Lang. Hear. Res. **49**, 138–149.
Sticht, T. G., and Gray, B. B. (**1969**). "The intelligibility of time compressed words as a function of age and hearing loss," J. Speech Hear. Res. **12**, 443–448.
Stine, E. A., and Wingfield, A. (**1987**). "Process and strategy in memory for speech among younger and older adults," Psychol. Aging **2**, 272–279.
Stine-Morrow, E. A., Soederberg-Miller, L. M., and Nevin, J. A. (**1999**). "The effects of context and feedback on age differences in spoken word recognition," J. Gerontol. B Psychol. Sci. Soc. Sci. **54**, 125–134.
Stuart, A., and Phillips, D. P. (**1996**). "Word recognition in continuous and broadband noise by young normal-hearing, older normal-hearing, and presbycusic listeners." Ear Hear. **17**, 478–489.
Summers, V., and Leek, M. R. (**1998**). "F0 processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss," J. Speech Lang. Hear. Res. **41**, 1294–1306.
Tremblay, K., Kraus, N., Carrell, T. D., and McGee, T. (**1997**). "Central auditory system plasticity: Generalization to novel stimuli following listening training," J. Acoust. Soc. Am. **102**, 3762–3773.
Tremblay, K., Kraus, N., McGee, T. J., Ponton, C. W., and Otis, B. (**2001**). "Central auditory plasticity: Changes in the N1-P2 complex following speech-sound training," Ear Hear. **22**, 79–90.
Trout, J. D. (**2005**). "Lexical boosting of noise-band speech in open- and closed-set formats," Speech Commun. **47**, 424–435.
Turner, C. W., Souza, P. E., and Forget, L. N. (**1995**). "Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners," J. Acoust. Soc. Am. **97**, 2568–2576.
van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (**1987**). "Speech waveform envelope cues for consonant recognition," J. Acoust. Soc. Am. **82**, 1152–1161.
van Tasell, D. J., Greenfield, D. G., Logemann, J. J., and Nelson, D. A. (**1992**). "Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of cochlear implants," J. Acoust. Soc. Am. **92**, 1247–1257.
Vaughan, N., and Letowski, T. (**1997**). "Effects of age, speech rate, and type of test on temporal auditory processing," J. Speech Lang. Hear. Res. **40**, 1192–1200.
Versfeld, N. J., and Dreschler, W. A. (**2002**). "The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners," J. Acoust. Soc. Am. **11**, 401–408.
Villaume, W. A., Brown, M. H., and Darling, R. (**1994**). "Presbycusis communication and older adults," in *Interpersonal Communication in Older Adulthood*, edited by M. L. Hummert, J. M. Weiman, and J. F. Nussbaum (Sage, Thousand Oaks, CA), pp. 83–106.
Vongpaisal, T., and Pichora-Fuller, M. K. (**2007**). "Effect of age on F_0 difference limen and concurrent vowel identification," J. Speech Lang. Hear. Res. **50**, 1139–1156.
Wingfield, A. (**1996**). "Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory," J. Am. Acad. Audiol **7**, 175–182.
Wingfield, A., Lindfield, K. C., and Goodglass, H. (**2000**). "Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition," J. Speech Lang. Hear. Res. **43**, 915–925.
Wingfield, A., Tun, P. A., Koh, C. K., and Rosen, M. J. (**1999**). "Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech," Psychol. Aging **14**, 380–389.
Wingfield, A., Tun, P. A., and McCoy, S. L. (**2005**). "Hearing loss in older adulthood: What it is and how it interacts with cognitive performance," Curr. Dir. Psychol. Sci. **14**, 144–148.
Wingfield, A., Wayland, S. C., and Stine, E. A. (**1992**). "Adult age differences in the use of prosody for syntactic parsing and recall of spoken sentences," J. Gerontol. **47**, 350–356.