Priming and sentence context support listening to noise-vocoded speech by younger and older adults

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

Citation: The Journal of the Acoustical Society of America **123**, 489 (2008); doi: 10.1121/1.2783762
View online: https://doi.org/10.1121/1.2783762
View Table of Contents: https://asa.scitation.org/toc/jas/123/1
Published by the Acoustical Society of America

Priming and sentence context support listening to noise-vocoded speech by younger and older adults

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

(Received 17 January 2007; revised 30 July 2007; accepted 22 August 2007)

Older adults are known to benefit from supportive context in order to compensate for age-related reductions in perceptual and cognitive processing, including when comprehending spoken language in adverse listening conditions. In the present study, we examine how younger and older adults benefit from two types of contextual support, predictability from sentence context and priming, when identifying target words in noise-vocoded sentences. In the first part of the experiment, benefit from context based on primarily semantic knowledge was evaluated by comparing the accuracy of identification of sentence-final target words that were either highly predictable or not predictable from the sentence context. In the second part of the experiment, benefit from priming was evaluated by comparing the accuracy of identification of target words when noise-vocoded sentences were either primed or not by the presentation of the sentence context without noise vocoding and with the target word replaced with white noise. Younger and older adults benefited from each type of supportive context, with the most benefit realized when both types were combined. Supportive context reduced the number of noise-vocoded bands needed for 50% word identification more for older adults than their younger counterparts.
 [DOI: 10.1121/1.2783762]

PACS number(s): 43.71.Lz, 43.71.Es, 43.71.Sy, 43.66.Sr [PEI] Pages: 489–499

I. INTRODUCTION

Older adults, even those with clinically normal audiometric thresholds in the speech range, have more difficulty than younger adults comprehending spoken language, especially in adverse listening conditions (for a review see CHABA, 1988; Pichora-Fuller and Souza, 2003). Even when older adults are able to correctly identify words heard in challenging listening conditions, there is evidence that they engage in more effortful processing than younger adults (for reviews see Pichora-Fuller, 2003; Wingfield and Tun, 2007). It has been suggested that older adults depend on supportive context to compensate for age-related reductions in perceptual and cognitive processing (e.g., Craik, 1983, 1986), including those that contribute to difficulties in comprehending spoken language (for reviews see Pichora-Fuller, 2003; Wingfield, 1996). Although both younger and older adults use various types of cues, including acoustic and lexical cues, to support listening to speech in challenging auditory scenes (Gallacher, 2005), older adults may even benefit more than younger adults from these various types of supportive context in challenging listening conditions (for reviews see Schneider and Pichora-Fuller, 2000; Wingfield and Tun, 2007). What is not yet fully understood is how supportive context from different sources, that is, acoustic cues from the speech signal and lexical cues from the syntactic or semantic context, interact during spoken language comprehension for younger and older listeners. Therefore, in the present study, we focus on the possible interactive use of two types of supportive context that are available when younger and older adults try to understand noise-vocoded speech.

A. Noise vocoding

Noise vocoding is a form of speech distortion that involves dividing a speech signal into specific frequency bands and, within each band, extracting the amplitude envelope and using it to modulate noise of the same bandwidth. This results in the fine structure of the signal being replaced with noise. This type of distortion minimizes the contribution of fine structure cues, and preserves the use of temporal amplitude-envelope cues. As the number of bands is increased, more band-specific envelope information becomes available. The intelligibility of noise-vocoded speech improves as the number of bands is increased (e.g., Loizou *et al.*, 1999; Shannon *et al.*, 1995).

There is compelling evidence that the comprehension of spoken language relies heavily on information that is carried by envelope cues (van Tasell *et al.*, 1992; Shannon *et al.*, 1995; Dorman and Loizou, 1998; Shannon, 2002). Specifically, envelope cues are one type of cue that carries supra-segmental information involved in lexical and syntactic processing (Schneider and Pichora-Fuller, 2001; Shannon, 2002; Greenberg, 1996; Rosen, 1992), with these cues contributing to speech prosody (Cutler *et al.*, 1997). Noise vocoding permits the investigation of the use of temporal amplitude-envelope cues relevant to supra-segmental speech processing while minimizing the contributions of fine structure cues that carry pitch-related supra-segmental information.

It is important to examine noise-vocoded speech comprehension for a number of reasons. First, given that we

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

know there are age-related declines in auditory temporal processing for cues relevant to segmental (Schneider *et al.*, 1994; Snell and Frisina, 2000; Pichora-Fuller *et al.*, 2006) and sub-segmental aspects of speech processing (Abel *et al.*, 1990; Summers and Leek, 1998; Alain *et al.*, 2001; Vongpaisal and Pichora-Fuller, 2007), it is important to discover if age-related differences also exist with regard to the envelope cues involved in supra-segmental speech processing. Second, noise vocoding provides an opportunity to examine how different types of contextual support may be used by younger and older listeners to compensate for auditory processing problems induced by challenging listening conditions (i.e., noise-vocoded speech).

In a recent companion study (Sheldon *et al.*, 2008), we used noise-vocoded speech to investigate possible age-related differences in the auditory processing of envelope cues. We established that noise-vocoding undermines word identification more for older than for younger listeners when target words are presented only once and with no feedback. Interestingly, these age-related differences in word identification were eliminated when older adults were able to benefit from feedback and/or from the summing of information over sequential presentations of the to-be-identified word, beginning with the word presented with one frequency band and incrementing the number of bands by one until it was correctly identified. Thus, noise vocoding provides an interesting type of signal degradation insofar as age-related perceptual differences in word identification seem to be effectively offset when feedback and/or summing of information is possible. One limitation of our earlier study is that there was no opportunity for the listeners to use different types of supportive context because the target words were spoken in the fixed carrier phrase "*Say the word*" (Sheldon *et al.*, 2008). Therefore, in the present experiment, we expand on the repertoire of types of supportive context to examine the resulting effect on lexical access.

B. Lexical access

Identifying a word involves establishing a mapping between the speech signal and the corresponding semantic meaning. This process is complex and intuitively integrative, involving information from both low-level and high-level processes. Accordingly, lexical access is affected by processes at these multiple levels. For instance, single-word lexical access can be supported via semantic priming (e.g., Meyer and Schvaneveldt, 1971) and phonological priming (e.g., Goldinger *et al.*, 1992). At the word level, semantic and phonological contexts have been found to each influence lexical selection; for example, Ferreira and Griffin (2003) showed that semantic priming and homophone priming could each improve lexical selection. Specific to speech, different types of context are known to mediate lexical selection during quiet listening conditions (e.g., Radeau *et al.*, 1998).

Less is known about the way in which different types of context combine to facilitate lexical access in challenging listening conditions and how such facilitation may change with age. It has been well established that in challenging listening conditions, speech can often be understood if there is sufficient semantic or linguistic contextual support available to provide information about the degraded signal (e.g., Kalikow *et al.*, 1977; Bilger *et al.*, 1984). Further, older adults are particularly adept at using context to compensate for difficulties hearing a degraded acoustic signal, presumably having developed expertise because typical everyday listening conditions are often more perceptually challenging for them than they are for younger adults (Perry and Wingfield, 1994; Pichora-Fuller *et al.*, 1995; Gordon-Salant and Fitzgibbons, 1997; Sommers and Danielson, 1999; Wingfield *et al.*, 2005).

The acoustic signal itself may also support spoken language comprehension by supplementing or augmenting the use of context based on semantic knowledge. Various studies have demonstrated that listeners can use phonological or prosodic information to direct attentional or top-down resources during spoken word recognition (Gow and Gordon, 1995; Marslen-Wilson and Tyler, 1980; Pitt and Samuel, 1990). Moreover, other situational cues, such as priming with a semantically related sentence (e.g., Gagné *et al.*, 2002), presenting visual speech for speech reading (e.g., Sumby and Pollack, 1954), presenting written text or clear speech as feedback (e.g., Davis *et al.*, 2005), spatially separating concurrent sounds (e.g., Freyman *et al.*, 1999, 2001; Li *et al.*, 2004), and increasing the pitch differences among simultaneous talkers (e.g., Mackersie and Prida, 2001) can all enhance speech intelligibility. These cues presumably enhance listening to the relevant structural patterns in the acoustic signal and may, in turn, enhance the effectiveness of semantic cues. In other words, information at the perceptual level can influence the deployment of semantic information for understanding degraded speech.

The interaction between acoustic and semantic level cues during lexical access has been examined by Connine *et al.* (1997). In their study, young, healthy participants performed a series of experiments that showed that phoneme detection reaction times are influenced by similarity to a carrier stimulus both in terms of form and meaning. Similarly, Andruski *et al.* (1994) demonstrated that the perceptual clarity of a word-initial phoneme is one determinant for the amount of semantic priming. Thus, these experiments demonstrate that the activation of the meaning of words is intimately tied to both the activation of the phonetic form of the word and to its lexical meaning.

The purpose of the present study is to investigate how different types of context interact to support word identification in younger and older adults. Specifically, we examine how the predictability of a word from sentence context, and priming with sentence content, combine to affect word identification in an adverse listening condition (noise-vocoded speech).

II. METHOD

A. Participants

Sixteen younger adults (mean age=20.6 years, s.d. =2.7, range=17–25 years) and sixteen older adults (mean age=67.6 years, s.d.=3.0, range=65–73 years) participated in this experiment. Both age groups had audiometric pure-

TABLE I. Mean (s.d.) of audiometric air-conducted pure-tone thresholds (dB HL) for test ears of younger and older participants.

	Frequency (kHz)							
	0.25	0.5	1	2	3	4	6	8
Younger participants ($N=16$)								
Mean	7.50	3.13	0.31	−2.18	−0.94	−0.63	3.13	1. 25
(s.d.)	(6.32)	(5.44)	(5.00)	(5.15)	(5.54)	(6.02)	(7.50)	(8.66)
Older participants ($N=16$)								
Mean	6.25	5.93	4.37	10.93	11.87	20.93	30.62	45.00
s.d.	(8.47)	(7.35)	(7.71)	(8.76)	(8.80)	(8.13)	(11.58)	(15.90)

tone air-conduction thresholds in the test ear that were considered to be clinically normal in the speech range (less than or equal to 25 dB HL from 0.25 to 3 kHz; see Table I). All of the participants were highly educated, with 15.1 (s.d.=2.7) and 13.7 (s.d.=2.3) mean years of education, respectively, for the younger and older groups; there was no significant difference in education between the age groups [$t(30)=1.48$, $p>0.10$]. All participants had learned English before the age of 5 years. To measure verbal knowledge, each participant completed the Mill-Hill Vocabulary Scale (Raven, 1965); the scores for both age groups indicated that all participants had good knowledge of the English language. The mean scores out of 20 for the younger and older groups were 12.5 (s.d. =2.2) and 14.8 (s.d.=2.2), respectively, with the older group significantly outperforming the younger group [$t(30)=0.30$, $p<0.05$]. All of the participants were paid volunteers recruited from the local community and none had previously heard noise-vocoded speech or the sentence materials that were used in the study. The participants completed Part A of the experiment in one session and returned within a week to complete Part B of the experiment.

B. Experiment: Part A

1. Materials

A digitized version of the sentences of the Speech Perception in Noise Test (SPIN-R; Bilger *et al.*, 1984) was used in the study. The SPIN-R materials consist of eight lists of 50 sentences per list. Each of the eight lists consists of 25 sentences in which the sentence-final word is predictable from the sentence context (e.g., "*Stir your coffee with a spoon*.") and 25 sentences in which the sentence-final word cannot be predicted from the sentence context (e.g., "*He would think about the rag*."). The high-context and low-context sentences are presented in a fixed pseudorandom order within each list.

For each list, the sentences were noise-vocoded into 16-band, 8-band, 4-band, and 2-band versions. To do so, we followed the procedure described in detail by Eisenberg *et al.* (2000). First, using the Goldwave digit audio editor, the stimuli were converted into binary files with a sampling rate of 20 kHz. Using MATLAB software, stimuli were then processed through a pre-emphasis filter [a high-pass first-order Butterworth infinite impulse response (IIR)] with a cut-off frequency of 1.2 kHz and an attenuation rate of −6 dB per octave. The signal was split into a varying number of frequency bands ($n=2,4,8,16$) 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 160 Hz). One minor difference between the present and earlier 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 kHz using a digital audio editor.

2. Procedure

During testing, the participant sat comfortably inside an International Acoustics Company (IAC) double-walled

TABLE II. Boundary frequencies (Hz) for the 2-, 4-, 8-, and 16-band noise-vocoded conditions.

2 band	4 band	8 band	16 band
300	300	300	300
1528	722	477	382
6000	1528	722	477
	3066	1061	590
	6000	1528	722
		2174	878
		3066	1061
		4298	1276
		6000	1528
			1825
			2174
			2584
			3066
			3632
			4298
			5080
			6000

sound-attenuating booth. The sound files were played by a Tucker Davis Technologies (TDT) System III monaurally to the participant's better ear over Sennheiser (model HD 265) headphones. The signal level was set at 70 dB SPL and was constant across conditions for both age groups.

To familiarize participants with noise-vocoded speech, each participant was first exposed to four lists of 50 noise-vocoded W-22 words (Martin and Pennington, 1971; Martin and Forbis, 1978; Penrod, 1994). The presentation of stimulus conditions was blocked, with the order of conditions progressing from easiest to hardest. First, participants heard a word list vocoded with 16 bands, followed by a word list in the 8-band condition, then a word list in the 4-band condition, and finally a word list in the 2-band condition. Words within a list were presented in random order and list order was counterbalanced across participants so that each list was presented four times in each order position for each group. For each word in each list, the participant was asked to identify the word. No feedback was provided either during the familiarization or during the experimental trials.

Following the familiarization with noise-vocoded words, participants were told that they would hear four lists of sentences and that they were to identify the last word of each sentence. One full SPIN-R list in the 16-band condition was presented, followed by another list in the 8-band condition, and then another list in the 4-band condition, and finally the last list in the 2-band noise-vocoded condition. Breaks were given as needed between lists. The list order was counterbalanced across participants so that each list was presented twice in each order position for each group. Guessing was strongly encouraged.

The experimenter used Sennheiser headphones (model HD 265) to listen to the participant's responses, which were scored immediately. If the experimenter was uncertain about any response, the participant was asked to repeat and spell the word aloud. Any phonemic difference between the response and the target word was marked as an error. No feedback was given after a response. All sessions were audio-taped to enable later verification of the responses recorded on the score sheet. The testing session lasted approximately 1–1.5 h.

C. Experiment: Part B

1. Materials

The same SPIN-R sentences were used in Part A and Part B of the experiment. Each participant heard all eight noise-vocoded SPIN-R lists, four lists in Part A and the other four lists in Part B of the experiment. No list was heard more than once by any participant. Part B differed from Part A because a priming utterance was presented prior to each sentence (following Experiment 2 of Freyman *et al.*, 2004). The priming utterance for each noise-vocoded sentence was constructed using the intact version of the sentence that had been used to produce the noise-vocoded sentence. For the prime, the initial portion of the intact sentence was presented, but the sentence-final word was replaced by a segment of white noise. The segment of noise used to replace the word was a randomly generated 700-ms white noise token that was scaled to have an average rms of 10 dB below the average rms of the noise-vocoded SPIN-R sentences. The wave form was edited to extract the sentence-final word from each sentence and to replace it with a noise segment such that the word and any coarticulatory cues were minimized in the priming utterance.

For presentation, the files were converted to single-channel .wav files and played monaurally to the participant's better ear over headphones. The signal level for the noise-vocoded sentences was set at 70 dB SPL and was constant across conditions for both age groups.

2. Procedure

During the testing session, the participant sat comfortably inside the double-walled sound-attenuating booth. Participants were told that they would hear four more lists of sentences and they were asked to listen to the prime (clear sentence context with the sentence-final word replaced with white noise) followed by the noise-vocoded sentence and then to identify the final word of the distorted sentence. First, they were presented one full SPIN-R list in the 16 band-processed condition, and then lists vocoded with 8, 4, and 2 bands. The order of the band conditions was fixed from easiest to hardest. Breaks were given as needed between lists. The list order was counterbalanced across participants so that each list was represented twice in each order position for each group. Guessing was strongly encouraged.

Participants' responses were scored immediately as in Part A. If the experimenter was uncertain about any response, the participant was asked to repeat and spell the word aloud. Any difference between the response and the target word was marked as an error. No feedback was given after a response. All sessions were audio-taped so that scoring could be confirmed as needed after the session. The testing session lasted approximately 1.5 h.

III. RESULTS

The percentage of sentence-final words correctly identified as a function of the number of frequency bands used in noise vocoding was calculated for each participant for each sentence type (low context or high context) in each experimental condition (with or without the priming utterance). We will refer to Part A of the experiment as the "without prime" experimental condition and Part B of the experiment as the "with prime" experimental condition.

Exponential functions of the form to follow were calculated for each of the two age groups to describe their performance in each of the four conditions; that is, with two levels of sentence context (low and high) and two levels of priming context (with and without prime):

$$y = c - e^{(a-bx)}.$$

Exponential functions were fit to the individual data, and to the average data for each age group. The functions describe the probability of correctly identifying sentence-final words (y) as a function of number of bands (x), subject to the restrictions that $a \geqslant 0$, and $c \leqslant 1$. Figure 1 plots the fitted functions for all of the participants in the low-context condi-

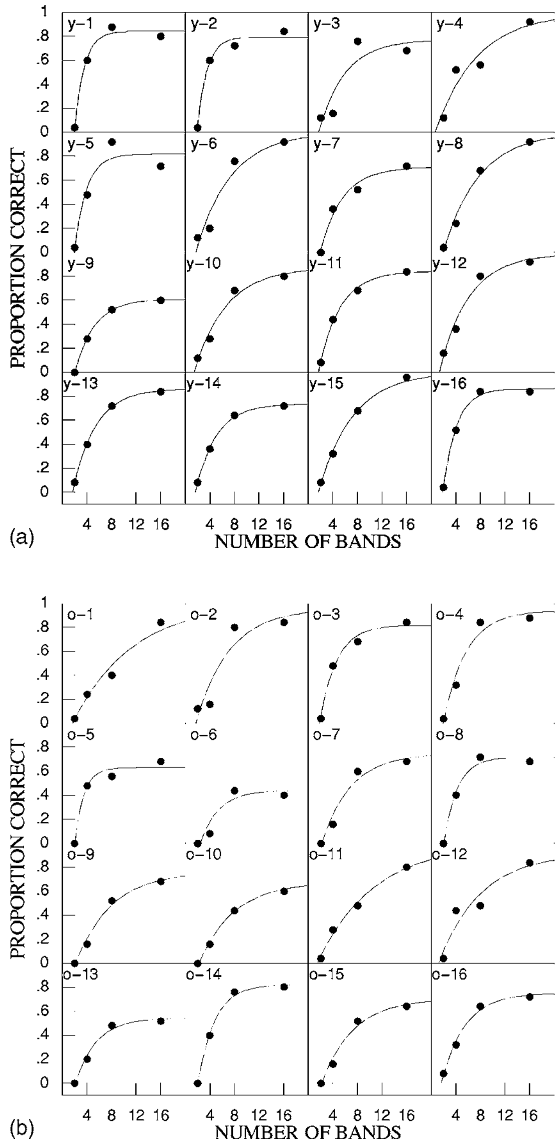

FIG. 1. (a) The proportion of words in low-context sentences that each younger participant (*y*) correctly identified, as a function of the number of bands used in noise vocoding. (b) The proportion of words in low-context sentences that each older participant (*o*) correctly identified, as a function of the number of bands used in noise vocoding.

tion with no prime. Figure 1 illustrates that the exponential function provides a good fit to the individual data.

The individual functions were used to estimate the band number that resulted in 50% of the target words being correctly identified in each of the four conditions (low context without prime; low context with prime; high context without prime; and high context with prime). For one older adult the 50% point could not be determined in the low-context condition without prime because for that participant the exponential function reached an asymptote that was less than 50%; for subsequent analyses, the data of this particular older participant were discarded and replaced with the corresponding mean of the data of the other older participants.

Figure 2 shows the raw mean band threshold for each of the four conditions for younger and older adults. As seen in Fig. 2, for both age groups, the mean band threshold is lowered by the presence of a prime, by context, and by the presence of both combined. Figure 2 also illustrates that in the low-context condition without prime older adults have a mean threshold value that is nearly two bands greater than the mean threshold value of younger adults. Figure 2 also indicates that the age-related difference in mean threshold values is greatly reduced when a prime is present (low context, with prime condition) and when contextual cues are present (high context, without prime). Furthermore, when both prime and high-context cues are available, there is no age-related difference in the mean threshold value. Thus, Fig. 2 suggests that the mean threshold values of older adults are reduced to a greater extent than those of younger adults by the presence of the prime, and also by the presence of high context, but to the same extent when both the prime and high context are available in combination.

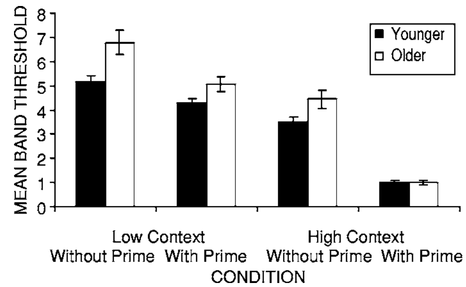

FIG. 2. The mean threshold values (the number of bands required to achieve 50% correct word identification) for older and younger participants in the four experimental conditions: Without prime, low context; with prime, low context; without prime, high context; with prime, high context. Standard error bars are shown.

The description of Fig. 2 was verified by an analysis of variance (ANOVA) with age (younger or older) as a between-subjects factor and prime condition (with or without prime) and sentence context (low context or high context) as within-subjects factors. There were significant main effects of age $[F(1,30)=7.88, p<0.01]$, prime $[F(1,30)=146.47, p<0.0001]$, and context $[F(1,30)=273.40, p<0.0001]$ on the number of noise-vocoded bands needed to achieve 50% word identification. There were also significant two-way interaction effects of prime and context $[F(1,30)=36.30, p<0.0001]$, age and prime $[F(1,30)=6.30, p<0.02]$, and age and context $[F(1,30)=4.46, p<0.05]$. The three-way interaction effect of age, prime, and context $[F(1,30)<1]$ was not significant.

The two-way interaction between prime and context indicates that the combined effects of priming and context are greater than the sum of their individual effects. To help interpret the pattern of two-way interaction effects between context and priming, age and priming, and age and context, we next examined the benefit that younger and older participants received from context in the conditions without and with prime. Figure 3 plots the mean proportion of instances in which the sentence-final word was correctly identified in low-context and high-context sentences as a function of the number of bands in two conditions (with prime and without prime) for the younger group and for the older group. The separations between the 50% points on the low-context and high-context functions are indicated by horizontal lines located at 0.5 on the abscissa.[1] Figure 3 suggests that there is a

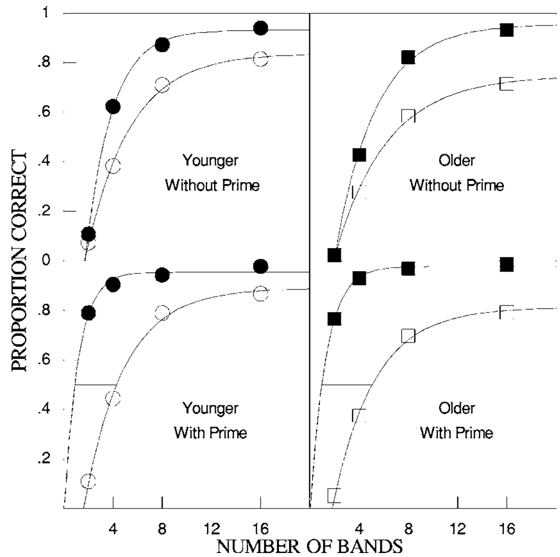

FIG. 3. The proportion of words correctly identified, averaged across participants, as a function of the number of bands for younger (circles) and older (squares) participants in the conditions without prime (Part A) and with prime (Part B). Closed symbols represent the conditions in which the sentence-final word was predictable from the sentence context; unfilled symbols represent the conditions in which the sentence-final word was not predictable from the sentence context.

larger difference in the number of bands required to achieve a threshold of 50% correct word identification between the low-context and high-context conditions in the condition with prime compared to the condition without prime. Moreover, the effect of priming on these thresholds appears to be about the same for both younger and older adults.

Although the extent of the age-related difference appears to be the same in the condition with prime as it is in the condition without prime, the threshold separation between the low-context and high-context functions appears to be larger in older adults than it is in younger adults in both the conditions with and without prime. To test this statistically, for the conditions with and without prime, we subtracted each individual's threshold for high-context sentences from his or her threshold for low-context sentences to obtain an estimate of the extent of the reduction in threshold due to context (the reduction in the number of bands needed to reach 50% correct).

Figure 4 plots the average reduction in threshold due to the addition of context for the two age groups in the conditions with and without prime (corresponding to the horizontal lines at 0.5 on the abscissa for each panel of Fig. 3). An ANOVA with age (younger or older) as a between-subjects factor and priming condition (with prime versus without prime) as a within-subjects factor confirmed that there was a significant main effect of age [$F(1,30)=4.53$, $p<0.05$], and a significant main effect of priming condition [$F(1,30)=36.27$, $p<0.0001$] on the benefit due to context, but there was no significant interaction between age and priming condition [$F(1,30)<1$)].

To complement Fig. 3, Fig. 5 plots the mean proportion of instances in which the sentence-final word was correctly identified in sentence contexts with and without prime as a function of the number of bands in two conditions (low context and high context) for the younger group and for the older group. In each panel of Fig. 5, the separation between the 50% points on the functions for the conditions with and without prime is indicated by horizontal lines located at 0.5 on the abscissa.[1] As seen in Fig. 5, the effects of priming are larger for high-context sentences than they are for low-context sentences by about the same amount in both younger and older adults. Figure 5 also shows that the effects of priming appear to be larger for older than they are for younger adults by about the same amount for both low-context and high-context sentences.

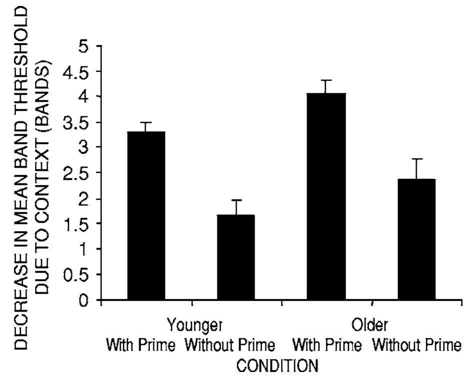

FIG. 4. The mean reduction in threshold (the number of bands required to achieve 50% correct word identification) due to sentence context for the four groups: younger adults, without prime; younger adults, with prime; older adults, without prime; and older adults, with prime. Standard error bars are shown.

To complement Fig. 4, Fig. 6 plots the average reduction in threshold due to priming for the two age groups in the high-context and low-context conditions (corresponding to

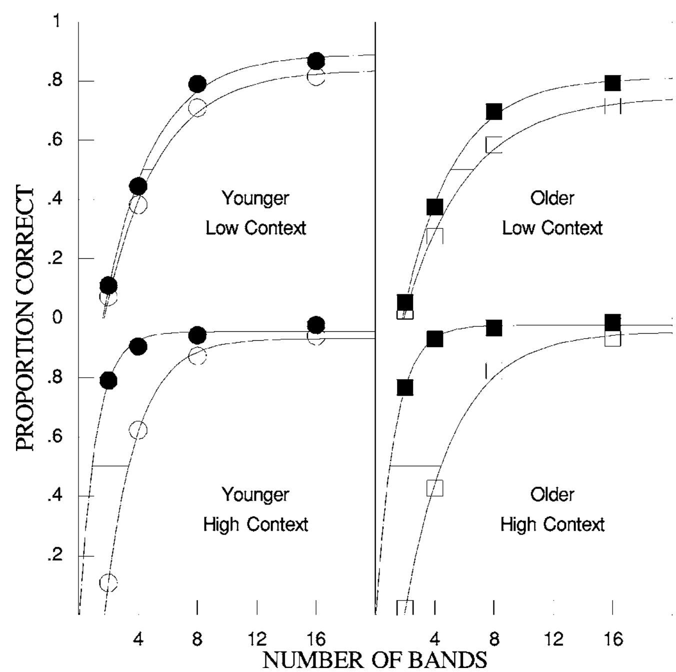

FIG. 5. The proportion of words correctly identified, averaged across participants, as a function of the number of bands for the conditions with prime (closed symbols) and without prime (open symbols) for younger adults (circles) and older adults (squares) when the sentence-final word was predictable from sentence context (high context) or when the sentence-final word was not predictable from sentence context (low context).

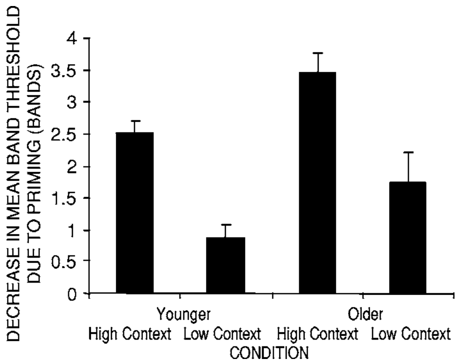

FIG. 6. The mean reduction in threshold (the number of bands required to achieve 50% correct word identification) due to the addition of the prime for the four groups: younger adults with low-context sentences; younger adults with high-context sentences, older adults with low-context sentences, and older adults with high context sentences. Standard error bars are shown.

the horizontal lines at 0.5 on the abscissa for each panel of Fig. 5). An ANOVA of the differences with age (younger and older) as a between-subjects factor and sentence context (low context or high context) as a within-subjects factor confirmed that there was a significant main effect of age [$F(1,30)=6.44$, $p>0.02$] and context [$F(1,30)=36.27$, $p<0.0001$], but no significant age by context interaction [$F(1,30)<1$].

IV. DISCUSSION

The primary goal of the current study was to explore age-related differences in the contributions of two types of supportive context: Support provided by priming with the undistorted presentation of all but the final word of the sentence, and support provided by the predictability of a sentence-final word from the sentence context in an adverse listening condition. We used a form of speech distortion, noise vocoding, in which the fine structure cues in the speech signal are systematically reduced to create different degrees of adversity. In general, the band threshold (number of bands corresponding to 50% correct target word identification) was higher for older than for younger adults, lower when the prime was present, and lower for high as opposed to low-context conditions. In addition, the band threshold in the condition combining high context and priming was 5 bands lower than in the condition with low context and without priming. This 5-band benefit from the combination of supports was bigger than the sum (3.3 bands) of the individual benefits of adding a prime alone (1.3 bands), or changing the context from low to high alone (2.0 bands).

The effects of context and priming were larger for older than for younger adults. Specifically, for younger adults, adding context reduced the threshold from 4.7 to 2.2 bands (2.5 bands); for older adults, it reduced the threshold from 5.9 to 2.7 bands (3.2 bands). Furthermore, for younger adults, priming reduced the threshold from 4.3 to 2.6 bands (1.7 bands); for older adults, it reduced the threshold from 5.6 to 3.0 bands (2.6 bands). However, these age-related differences disappear if the effects of context and priming are considered in terms of the percentage decrease in threshold.[2] Specifically, the percentage decrease when going from low to high context was 48% for younger and 46% for older adults. Furthermore, going from conditions without prime to conditions with prime reduced thresholds by 61% and 54% for younger and older adults, respectively. Hence, depending on whether or not the threshold changes are expressed as differences or percentages, we would conclude either that older adults benefit more from priming and context that do younger adults (thresholds changes expressed as differences), or that younger and older adults benefit equally from these two factors (threshold changes expressed as percentages). In either case, we can conclude that older adults benefit at least as much as do younger adults from these two types of support.

It seems unlikely that the age differences found in the present study can be attributed to age-related differences in audiometric thresholds. All younger and older adults had clinically normal audiometric thresholds for frequencies from 0.25 to 3 kHz, although the pure-tone average (0.5, 1, and 2 kHz) of the older group was about 6 dB higher than that of the younger group, and there were larger differences at higher frequencies. Nevertheless, it is not obvious how small differences in hearing threshold could account for the pattern of results related to the number of bands used in noise vocoding the sentences. In particular, increasing the number of bands used to vocode the sentences provided the listeners with more frequency-specific envelope information, but the frequency range and the audibility of the noise-vocoded sentences remained constant across the different band conditions. Furthermore, all stimuli were presented in quiet so age-related differences related to signal-to-noise ratio cannot explain the pattern of findings. Importantly, our finding of an age effect in the present study is consistent with our finding of age-related differences in Experiment 2 in our previous study (Sheldon *et al.*, 2008), and it is also 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.

Considering other possible factors associated with age, the participants did not differ in education level. However, older adults outperformed younger adults on the Mill-Hill vocabulary score (Raven, 1965). Hence, the possibility that the greater benefit from support realized by the older adults might be related to their superior lexical knowledge cannot be ruled out.

The results of the study will be discussed with an emphasis on age-related differences in the auditory processing of temporal amplitude-envelope cues relevant to suprasegmental speech processing and the use of the two types of support and how they may combine during spoken language comprehension.

A. Age-related differences in auditory temporal processing

The word identification scores in the low-context condition without prime (Part A) indicate that older adults are poorer than younger adults at processing envelope cues when little contextual support is available. The present results for sentence-final word identification in the low-context condition without prime are consistent with the results of our earlier study in which age-related deficits were found in the use of envelope cues to identify noise-vocoded monosyllable words presented with a fixed carrier phrase, at least when there was no feedback or opportunity to sum information from sequential presentations of the same word (Sheldon *et al.*, 2008). Nevertheless, for the participants in the present experiment, we found no significant correlation between their word identification performance in the familiarization task and their performance in any of the four experimental conditions (low context with or without prime, high context with or without prime).

Together with prior research, the present findings provide converging evidence that aging is associated with deficits at multiple levels of auditory temporal processing that contribute to different levels of speech processing: the subsegmental level (e.g., Abel *et al.*, 1990; Summers and Leek, 1998; Alain *et al.*, 2001; Vongpaisal and Pichora-Fuller, 2007), the segmental level (Schneider *et al.*, 1994; Snell and Frisina, 2000; Pichora-Fuller *et al.*, 2006), and the suprasegmental level (Sheldon *et al.*, 2008).

B. Sentence context supports spoken language comprehension

When younger and older participants listened to SPIN-R sentences in the four band-processed conditions, whether they heard a prime (Part B) or did not (Part A), word identification performance was better in high-context sentences than in low-context sentences for both age groups. However, the difference due to the addition of this type of context was greater for older than for younger listeners. In other words, older adults derived more benefit from sentence context than did younger adults when benefit was measured in terms of the difference in the number of bands required for 50%-correct identification. The present results are consistent with previous studies that have shown that older adults are the same (Dubno *et al.*, 2000) or even better than younger adults at using semantic context than younger adults to boost word identification in adverse situations where listening (or reading) is effortful (for reviews see Schneider, 2001; Pichora-Fuller, 2003). A general finding has been that older adults require a higher signal-to-noise ratio to achieve the same word identification score as younger adults when no supportive context is available. Because the signal-to-noise conditions typical in everyday life would often be challenging for older listeners even though these conditions would not be particularly challenging for younger listeners, one reasonable possibility is that older adults have much more practice at using context in a compensatory fashion (Pichora-Fuller *et al.*, 1995). It is important to note that the older participants in the present study scored higher than did the younger participants on the vocabulary test. The superior vocabulary of the older participants in the present study is consistent with the possibility that the larger lexicons and greater lexical familiarity among older adults could explain an age-related benefit from sentence context associated with better semantic knowledge (Wingfield *et al.*, 2005).

C. Priming directs auditory processing

The presence of the prime also decreased the mean band threshold for both age groups. Even when semantic and linguistic cues were minimal, as in the low-context sentences, there was still an increase due to priming in the word identification scores for both age groups. Because the comprehension of noise-vocoded speech is primarily based on information carried by the envelope, it is possible that listeners attend to the envelope cues of the intact prime to facilitate comprehension of the vocoded sentence. That is, the prime may direct a listener's attention to relevant aspects of structure of the speech signal, thereby contributing to improved sentence-final word identification. Support for this explanation comes from studies that stress the importance to comprehension of the supra-segmental speech information provided by the envelope (Martin, 1979; for a review, see Cutler *et al.*, 1997). Specifically, envelope cues preserve amplitude modulations that may provide cues to the beginnings and endings of words that are needed for segmenting running speech (Sanders *et al.*, 2002; Sanders and Neville, 2003a, b). It also seems that the prime may offer an opportunity for perceptual learning that could result in improved understanding of noise-vocoded speech. This is suggested by the finding that when clear speech or written text was provided following the presentation of a noise-vocoded utterance, word identification improved for subsequently presented novel utterances (Davis *et al.*, 2005).

The presence of the prime was of greater benefit to older than younger adults in the low-context conditions when the advantage is expressed in terms of differences in number of bands required for 50%-correct identification. To our knowledge, an age-related advantage from this sort of priming has not been reported previously. We offer two possible explanations for our observations.

1. Compensation. One possibility is that the age-related difference in ability to benefit from supportive context reflects how older adults compensate by engaging top-down processing to achieve the same performance as younger adults on various perceptual and cognitive tasks (for a review see Pichora-Fuller and Singh, 2006). The possibility that older adults engage in different and possibly more effortful types of information processing to compensate for deficits in auditory processing is consistent with the more general observation that older adults use more brain regions and show a reduction in hemisphericity of activation on a wide range of perceptual and cognitive tasks. This age-related difference in activation is thought to reflect compensatory adaptations (e.g., Grady, 2000; Cabeza, 2002; Reuter-Lorenz, 2002). The age-related difference in benefit from the prime in the low-context sentence conditions may reflect the greater compensatory use of supportive context by older adults in challenging listening conditions such as those that they frequently encounter in everyday life. We speculate that older adults are simply more skilled at using information provided by the prime to direct attention to the envelope cues in the vocoded speech that facilitate word segmentation and phonological processing, particularly in low-context sentences where other types of contextual support are not readily available.

2. Summing information. Another explanation is that there is an age-related difference in the processing strategies used when the prime is available. In our earlier study, older listeners were able to match the word identification accuracy of younger listeners by summing information across a sequence of versions of the same target word that were incremented progressively in the number of bands used to vocode the speech (Sheldon *et al.*, 2008). In the current study, the listener may be able to sum information by comparing the clear sentence context of the prime and the vocoded sentence context. It could be that older adults are particularly adept at summing knowledge to help decipher a target word.

D. Priming facilitates benefit from sentence context

Not surprisingly, when both types of support, priming and high sentence context, are available, listeners reach ceiling performance and no significant age-related differences are observed. The benefit derived when these two types of support are combined is greater than the sum of their independent contributions. In the high-context condition with prime, both younger and older listeners correctly identified over half of the words even in the 2-band condition, the lowest band condition tested. This high level of performance indicates that both age groups successfully use auditory verbal closure such that contextual information facilitates identification of the degraded target word (for a review see Elliott, 1995). The prime facilitates the use of sentence context because it provides an undistorted presentation of the sentence context. However, even when the sentence context is low, the prime provides information that may help the listener to narrow in on what to listen for and when to listen in the noise-vocoded sentence carrying the target word (Freyman *et al.*, 2004). It seems that listeners benefit from knowledge of the acoustical structure as well as semantic knowledge of the sentence context.

V. CONCLUSIONS

Envelope cues that provide supra-segmental speech information are integral to spoken language comprehension. When such cues are degraded, as they are when speech is noise vocoded, younger and older listeners compensate by relying to a greater extent on the available supportive context. Five conclusions can be drawn from this study regarding the use of such support when speech is noise vocoded:

(1) Words that are highly predictable from sentence context are better comprehended than words that are not predictable from sentence context.
(2) A prime, in the form of undistorted presentation of the sentence context, can facilitate comprehension, even in the absence of high-context semantic information, by directing a listener to relevant acoustic properties of the speech signal.
(3) When the prime is presented for a high-context sentence, the prime also facilitates benefit from the semantic context, thus demonstrating the interactivity of the two types of cues.
(4) Older adults do not identify words as well as younger adults when envelope cues are reduced by noise vocoding. Conversely, older adults need more bands than younger adults to achieve the same level of accuracy in identifying the final words of noise-vocoded sentences.
(5) Older adults benefit more than younger adults from sentence context and the prime; however, both age groups receive equal benefit when these two types of context interact.

ACKNOWLEDGMENTS

The authors are grateful to Ewen MacDonald for creating the noise-vocoding software. This research was funded by the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada (NSERC). These experiments were conducted as part of the Master's Thesis of S.S.

[1]Note that to find the 50% correct points on the functions for the high-context sentences with prime shown in Figs. 3 and 5, it was necessary to extrapolate from the data. For both age groups, the mean percent correct score exceeded 70% when high-context sentences were presented with a prime in the 2-band noise-vocoding conditions. To find the 50% correct point, we extrapolated using the conservative assumption that both groups would score 0% correct if no stimulus were presented (0 band condition). Given that the sentence-final words in the high-context sentences of the SPIN test were designed to be highly predictable from preceding sentence context, a more realistic and less conservative assumption would have been a minimum score of 70% based on the paper and pencil test of auditory verbal closure that was conducted during the development of the original SPIN test (Kalikow *et al.*, 1977; see also Bilger *et al.*, 1984; Elliott, 1995). Had we used a less conservative method to estimate the 50% point, or had we used a percentage correct higher than 50% to compare conditions, then the difference due to context (Fig. 3) and the difference due to priming (Fig. 5) would have been larger for both age groups, but the differences for the older group would have remained larger than those for the younger group.

[2]To confirm the lack of interaction when changes in the 50% correct-identification thresholds are expressed as percentages, we conducted an ANOVA on the log band thresholds with age (younger or older) as a between-subjects factor and priming condition (with or without prime) and sentence context (low context or high context) as within-subjects factors. Note that when the 50% band threshold values are converted to logarithms, equal ratios between any two pairs of threshold values on the original measures correspond to equal intervals on the log-transformed values. All of the main effects remained significant: age $[F(1,30)=4.66, p<0.05]$, prime $[F(1,30)=516.75, p<0.0001]$, and context $[F(1,30)=437.43, p<0.0001]$ on the number of noise-vocoded log bands needed to achieve 50% word identification. The significant two-way interaction of prime and context also remained significant, $[F(1,30)=311.00, p<0.0001]$; however, the two-way interactions with age were not significant: age and prime $[F(1,30)=3.25, p>0.81]$, and age and context $[F(1,30)<1]$. The three-way interaction effect of age, prime, and context $[F(1,30)<1]$ was again not significant.

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.

Andruski, J. E., Blumstein, S. E., and Burton, M. (**1994**). "The effect of subphonetic differences in lexical access," Cognition **52**, 163–187.

Bilger, R. C., Nuetzel, M. J., Rabinowitz, W. M., and Rzeczkowski, C. (**1984**). "Standardization of a test of speech perception in noise," J. Speech Hear. Res. **27**, 32–48.

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

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.

Connine, C. M., Titone, D., Deelman, T., and Blasko, D. (**1997**). “Similarity mapping in spoken word recognition,” J. Mem. Lang. **37**, 463–480.

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, Amsterdam), pp. 499–522.

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

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.

Dorman, M. F., and Loizou, P. C. (**1998**). “Identification of consonant and vowels by cochlear implant patients using a 6-channel continuous interleaved sampling processor and by normal hearing subjects using simulations processors with two to nine channels,” Ear Hear. **19**, 162–166.

Dubno, J. R., Ahlstrom, J. B., and Horwitz, A. R. (**2000**). “Use of context by young and aged adults with normal hearing,” J. Acoust. Soc. Am. **107**, 538–546.

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.

Elliott, L. L. (**1995**). “Verbal auditory closure and the Speech Perception in Noise (SPIN) Test,” J. Speech Hear. Res. **38**, 1363–1376.

Ferreira, V. S., and Griffin, Z. M. (**2003**). “Phonological influences on lexical (mis)selection,” Psychol. Sci. **14**, 86–90.

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (**2001**). “Spatial release from informational masking in speech recognition,” J. Acoust. Soc. Am. **109**, 2112–2122.

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (**2004**). “Effect of number of masking talkers and auditory priming on informational masking in speech recognition,” J. Acoust. Soc. Am. **115**, 2246–2256.

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (**1999**). “The role of perceived spatial separation in the unmasking of speech,” J. Acoust. Soc. Am. **106**, 3578–3588.

Gagné, J.-P., Rochette, A.-J., and Charest, M. (**2002**). “Auditory, visual, and audiovisual clear speech,” Speech Commun. **37**, 213–230.

Gallacher, J. (**2004**). “Hearing, cognitive impairment and aging: A critical review,” Reviews in Clinical Gerontology **14**, 199–209.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., and Marcario, J. K. (**1992**). “Form-based priming in spoken word recognition; the roles of competition and bias,” J. Exp. Psychol. Learn. Mem. Cogn. **18**, 1211–1238.

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.

Gow, D. W., and Gordon, P. C. (**1995**). “Lexical and prelexical influences on word segmentation: Evidence from priming,” J. Exp. Psychol. Hum. Percept. Perform. **21**, 344–359.

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.

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

Kalikow, D. N., Stevens, K. N., and Elliott, L. L. (**1977**). “Development of a test of speech intelligibility in noise using sentence material with controlled word predictability,” J. Acoust. Soc. Am. **61**, 1337–1351.

Li, L., Daneman, M., Qi, J., and Schneider, B. A. (**2004**). “Does the information content of an irrelevant source differentially affect speech recognition in younger and older adults?,” J. Exp. Psychol. Hum. Percept. Perform. **30**, 1077–1091.

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.

Mackersie, C. L., and Prida, T. L. (**2001**). “The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss,” J. Speech Lang. Hear. Res. **44**, 19–28.

Marslen-Wilson, W. D., and Tyler, L. K. (**1980**). “The temporal structure of spoken language understanding,” Cognition **8**, 1–71.

Martin, F. N., and Forbis, N. R. (**1978**). “The present status of audiometric practice: A follow-up study,” ASHA **20**, 531–541.

Martin, F. N., and Pennington, C. D. (**1971**). “Current trends in audiometric practices,” ASHA **13**, 671–677.

Martin, J. G. (**1979**). “Rhythmic and segmental perception are not independent,” J. Acoust. Soc. Am. **65**, 1286–1297.

Meyer, D. E., and Schvaneveldt, R. W. (**1971**). “Facilitation in recognizing pairs of words; evidence of a dependence between retrieval operations,” J. Exp. Psychol. **90**, 227–234.

Penrod, J. P. (**1994**). “Speech threshold and word recognition/discrimination testing,” in *Handbook of Clinical Audiology*, 4th ed., edited by J. Katz (Williams and Wilkins, Baltimore, MD), pp. 147–164.

Perry, A. R., and Wingfield, A. (**1994**). “Contextual encoding by young and elderly adults as revealed by cued and free recall,” Aging, Neuro., and Cogn. **1**, 120–139.

Pichora-Fuller, M. K. (**2003**). “Cognitive aging and auditory information processing.” Int. J. Audiol., Suppl. 2, **42**, S26–S32.

Pichora-Fuller, M. K., Schneider, B. A., Benson, N., Hamstra, S., and Storzer, E. (**2006**). “Effect of age on gap detection in speech and nonspeech stimuli varying in marker duration and spectral symmetry,” J. Acoust. Soc. Am. **119**, 1143–1155.

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 Singh, G. (**2006**). “Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation,” Trends Amplif. **10**, 29–59.

Pichora-Fuller, M. K., and Souza, P. (**2003**). “Effects of aging on auditory processing of speech,” Int. J. Audiol., Suppl. @, Suppl 2 **42**, S11–S16.

Pitt, M. A., and Samuel, A. G. (**1990**). “The use of rhythm in attending to speech,” J. Exp. Psychol. Hum. Percept. Perform. **16**, 564–573.

Radeau, M., Besson, M., Fonteneau, E., and Castro, S. L. (**1998**). “Semantic, repetition and rime priming between spoken words: Behavioral and electrophysiological evidence,” Biol. Psychol. **48**, 183–204.

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.

Sanders, L. D., and Neville, H. J. (**2003a**). “An ERP study of continuous speech processing. II. Segmentation, semantics, syntax in non-native speakers,” Brain Res. Cognit. Brain Res. **15**, 214–227.

Sanders, L. D., and Neville, H. J. (**2003b**). “An ERP study of continuous speech processing. I. Segmentation, semantics, syntax in native speakers,” Brain Res. Cognit. Brain Res. **15**, 228–240.

Sanders, L. D., Newport, E. L., and Neville, H. J. (**2002**). “Segmenting nonsense: An event-related potential index of perceived onsets in continuous speech,” Nature (London) **5**, 700–703.

Schneider, B. A. (**2001**). “Sensation, cognition, and levels of processing in aging,” in *Perspectives on Human Memory and Cognitive Aging: Essays in Honour of Fergus Craik*, edited by M. Naveh-Benjamin, M. Moscovitch, and H. L. Roediger III (Psychology Press, New York), pp. 298–314.

Schneider, B. A., and Pichora-Fuller, M. K. (**2000**). “Implications of perceptual deterioration for cognitive aging research.” in *Handbook of Aging and Cognition*, 2nd, ed., edited by F. I. M. Craik and T. A. Salthouse (Erlbaum, Mahwah, NJ), pp. 155–220.

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., Kowalchuk, 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,” Am. J. of Audiol. **11**, 124–127.

Shannon, R. V., Zeng, F., Kamath, V., and Wygonski, J. (**1995**). “Speech recognition with primarily temporal cues,” Science **270**, 303–304.

Sheldon, S., Pichora-Fuller, M. K., Schneider, B. A. (**2008**). “Effect of age, presentation method, and training on identification of noise-vocoded words,” J. Acoust. Soc. Am. **123**, ■.

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., and Danielson, S. M. (**1999**). “Inhibitory processes and spoken word recognition in young and old adults: The interaction of lexical competition and semantic context,” Psychol. Aging **14**, 458–472.
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.
Sumby, W. H., and Pollack, I. (**1954**). “Visual contribution to speech intelligibility in noise,” J. Acoust. Soc. Am. **26**, 212–215.
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.
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.
Vongpaisal, T., and Pichora-Fuller, M. K. (**2007**). “Effect of age on F0 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., and Tun, P. A. (**2007**). “Cognitive supports and cognitive constraints on comprehension of spoken language,” J. Am. Acad. Audiol **18**.
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.