Word Recognition for Temporally and Spectrally Distorted Materials: The Effects of Age and Hearing Loss

Sherri L. Smith,[1] Margaret Kathleen Pichora-Fuller,[2] Richard H. Wilson,[1] and Ewen N. MacDonald[3]

Objectives: The purpose of Experiment 1 was to measure word recognition in younger adults with normal hearing when speech or babble was temporally or spectrally distorted. In Experiment 2, older listeners with near-normal hearing and with hearing loss (for pure tones) were tested to evaluate their susceptibility to changes in speech level and distortion types. The results across groups and listening conditions were compared to assess the extent to which the effects of the distortions on word recognition resembled the effects of age-related differences in auditory processing or pure-tone hearing loss.

Design: In Experiment 1, word recognition was measured in 16 younger adults with normal hearing using Northwestern University Auditory Test No. 6 words in quiet and the Words-in-Noise test distorted by temporal jittering, spectral smearing, or combined jittering and smearing. Another 16 younger adults were evaluated in four conditions using the Words-in-Noise test in combinations of unaltered or jittered speech and unaltered or jittered babble. In Experiment 2, word recognition in quiet and in babble was measured in 72 older adults with near-normal hearing and 72 older adults with hearing loss in four conditions: unaltered, jittered, smeared, and combined jittering and smearing.

Results: For the listeners in Experiment 1, word recognition was poorer in the distorted conditions compared with the unaltered condition. The signal to noise ratio at 50% correct word recognition was 4.6 dB for the unaltered condition, 6.3 dB for the jittered, 6.8 dB for the smeared, 6.9 dB for the double-jitter, and 8.2 dB for the combined jitter-smear conditions. Jittering both the babble and speech signals did not significantly reduce performance compared with jittering only the speech. In Experiment 2, the older listeners with near-normal hearing and hearing loss performed best in the unaltered condition, followed by the jitter and smear conditions, with the poorest performance in the combined jitter-smear condition in both quiet and noise. Overall, listeners with near-normal hearing performed better than listeners with hearing loss by ~30% in quiet and ~6 dB in noise. In the quiet distorted conditions, when the level of the speech was increased, performance improved for the hearing loss group, but decreased for the older group with near-normal hearing. Recognition performance of younger listeners in the jitter-smear condition and the performance of older listeners with near-normal hearing in the unaltered conditions were similar. Likewise, the performance of older listeners with near-normal hearing in the jitter-smear condition and the performance of older listeners with hearing loss in the unaltered conditions were similar.

Conclusions: The present experiments advance our understanding regarding how spectral or temporal distortions of the fine structure of speech affect word recognition in older listeners with and without clinically significant hearing loss. The Speech Intelligibility Index was able to predict group differences, but not the effects of distortion. Individual differences in performance were similar across all distortion conditions with both age and hearing loss being implicated. The speech materials needed to be both spectrally and temporally distorted to mimic the effects of age-related differences in auditory processing and hearing loss.

(Ear & Hearing 2012;33;349–366)

[1]Research and Development Service, Department of Veterans Affairs, James H. Quillen Veterans Affairs Medical Center, Mountain Home, Tennessee and Department of Audiology and Speech-Language Pathology, East Tennessee State University, Johnson City, Tennessee; [2]Department of Psychology, University of Toronto, Mississauga, Ontario and Toronto Rehabilitation Institute, Toronto, Ontario, Canada and Department of Behavioural Sciences and Learning, Linköping University, Sweden; and [3]Department of Psychology, Queen's University, Kingston, Ontario, Canada and Centre for Applied Hearing Research, Department of Electrical Engineering, Technical University of Denmark, Lyngby, Denmark

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

INTRODUCTION

Older listeners have reduced abilities to understand speech compared with younger listeners with normal-hearing sensitivity, especially when background noise is present. Recent research has focused on differentiating the subtypes of presbycusis and the roles of spectral resolution or temporal processing in word recognition by listeners with sensorineural hearing loss (SNHL), especially when there is background noise or competing speech (for reviews, see Moore 2008; Pichora-Fuller & MacDonald 2008). In particular, researchers have been interested in how subtypes of presbycusis might differ physiologically (Schmiedt 2010) and behaviorally (Fitzgibbons & Gordon-Salant 2010) from each other and from types of SNHL typical in younger adults.

SNHL was characterized by Carhart (1951) as having two components, one relating to the loss of acuity and the other relating to the loss of clarity. Plomp (1978) referred to these two components as attenuation and distortion, respectively. Speech perception difficulties in older listeners with SNHL have been attributed in part to the elevation of thresholds (i.e., reduced audibility) and to internal distortions in auditory processing arising from various types of damage to the cochlea or the neural pathways (e.g., Bocca & Calearo 1963; Frisina & Frisina 1997; Gates & Mills 2005; Mills et al. 2006; Humes & Dubno 2010). Such internal distortions may alter suprathreshold spectral or temporal processing of fine structure cues (defined as rapid fluctuations in the time waveform; e.g., Rosen 1992; Moore 2008) and may depend also on the etiology of the hearing loss. Furthermore, distortions arising from neural damage may occur in the absence of significant threshold elevation or abnormalities in otoacoustic emissions (Willott 1991; Kujawa & Liberman 2009). The type and amount of internal distortion associated with aging or SNHL and the effect that such distortion has on speech perception, however, continues to be active areas of investigation.

0196/0202/12/3303-0349/0 • Ear & Hearing • Copyright © 2012 by Lippincott Williams & Wilkins • Printed in the U.S.A.

Modeling SNHL

Models have been developed to predict the intelligibility of speech for a listener with a given audiometric hearing loss. One such model commonly used today is the Speech Intelligibility Index (SII; American National Standards Institute [ANSI] S3.5-R 1997), which largely extends upon early works related to the Articulation Index (AI; first proposed by French & Steinberg 1947). To predict speech intelligibility, the SII uses information about variables related to the speech spectrum level, the noise spectrum level, the relative band importance functions of the speech materials, and the pure-tone thresholds of the listener group. Even though the SII expanded upon the AI with a number of changes, including the use of frequency bands of different weights, data show that discrepancies remain between SII predictions and the observed speech performances of many listeners with hearing loss. Such discrepancies suggest that the effects of hearing loss on speech intelligibility are not attributable solely to audibility issues as indexed by audiometric threshold elevations (Hargus & Gordon-Salant 1995; Ching et al. 1998; Hogan & Turner 1998; Ching et al. 2001; Humes & Dubno 2010), but that they also reflect changes in suprathreshold processing or the clarity or distortion components of hearing loss as described by Carhart (1951) and Plomp (1978), respectively. Changes in suprathreshold processing are not necessarily well predicted by audiometric thresholds and may vary considerably across individuals, especially older individuals (Plomp 1986).

The AI and SII models are accurate in predicting the recognition performance of many listeners with mild to moderate pure-tone hearing losses who have good recognition performance, presumably because these listeners are mostly affected by reduced speech audibility (e.g., Kamm et al. 1985). The models, however, do not accurately reflect the distortion component of SNHL, such as reductions in spectral and temporal resolution, or other factors thought to affect listeners with more severe degrees of hearing losses (Ching et al. 2001) or those who exhibit unusually poor speech intelligibility. In addition, the models have not accounted successfully for performance degradations owing to the upward spread of masking or decreases in speech performance at high presentation levels (Kamm et al. 1985; Studebaker et al. 1999; Ching et al. 2001; Kates & Arehart 2005). Overall, the SII can provide a useful and reasonably accurate index of how reduced audibility will likely affect word recognition and it has been successful in accounting for much of the variance in the performance of older adults with various degrees of high-frequency hearing loss when they listen to speech in quiet or steady state background noise; however, there remain limitations in how well the SII can predict performance for distorted speech presented with fluctuating competing noise to listeners with deficits in suprathreshold auditory processing that may be independent of audiometric threshold elevations (Humes & Dubno 2010). The shortcomings of the models are likely to be relevant when considering the effects on speech intelligibility of suprathreshold distortions that may arise from the auditory temporal processing problems frequently found in older listeners who have near-normal audiometric thresholds (Fitzgibbons & Gordon-Salant 2010).

Simulating SNHL

Simulations may provide insights into how the so-called audibility and distortion factors contribute to word recognition. Applying distortions to speech materials does not guarantee that anatomical and physiological changes independent of or secondary to hearing loss or age are being simulated accurately in terms of reproducing in listeners with normal hearing the experiences of those with actual hearing loss, but investigators have been successful in equating performance between listener groups with given types of simulations and with given sets of stimuli (e.g., Baer & Moore 1993; Pichora-Fuller et al. 2007). Thus, at least to some degree, this approach advances our knowledge of how specific manipulations of signal properties can affect performance.

In listeners with outer hair cell damage, reduced frequency selectivity corresponds to broadened tuning curves or widened auditory filters, with one consequence being that the frequency components of speech are more difficult to resolve and masking may be greater than would be the case in listeners with normal hearing (e.g., Moore et al. 1992). It is important to note that a loss of frequency selectivity in listeners with SNHL is thought to contribute to their speech perception difficulties, particularly in competing noise. Baer and Moore (1993) evaluated the extent to which sentence recognition in quiet and in noise was affected by simulations of reduced frequency selectivity. Younger listeners with normal-hearing sensitivity were presented sentences in which the spectra were "smeared" to mimic the reduced frequency selectivity of listeners who have moderate or severe cochlear hearing loss. Baer and Moore showed that sentence recognition in quiet was unaffected by spectral smearing. Sentence recognition in noise, however, was reduced significantly, particularly for greater degrees of spectral smearing at poorer signal to noise ratios (SNRs). Although spectral smearing had no significant effect on sentence recognition in quiet, Boothroyd et al. (1996) did find that spectral smearing affected both phoneme- and word-recognition performances in quiet. Similar to the findings of Baer and Moore for sentences, Boothroyd et al. found that phoneme recognition was affected by spectral smearing and that it was exacerbated with the addition of noise. The results of these studies suggest that frequency selectivity is important for speech perception in noise, but may be less important for speech perception in quiet, at least with sentences rich in linguistic structure and context.

Although the importance of spectral cues for speech perception has been studied extensively, the importance of temporal cues has not been examined until relatively recently. Listeners with cochlear hearing loss exhibit reduced temporal resolution, which may contribute to the speech perception difficulties often reported by listeners with hearing loss (Lorenzi & Moore 2008; Reed et al. 2009). Furthermore, older listeners, even those with minimal pure-tone threshold elevations, demonstrate reduced temporal processing abilities and speech perception difficulties (Gordon-Salant & Fitzgibbons 1999; Pichora-Fuller & Souza 2003). Thus, there is a strong suggestion of a connection between temporal processing deficits and speech perception.

Temporal envelope and fine structure cues serve various roles in speech perception (for reviews see Rosen 1992; Greenberg 1996). It is important to note that age-related differences have been found in the temporal processing of envelope and fine structure cues involved in the processing of suprasegmental, segmental, and voice information (for reviews

see Schneider & Pichora-Fuller 2001; Pichora-Fuller & MacDonald 2008). The pattern of fluctuations in the temporal envelope of a speech wave form provides suprasegmental prosodic cues for rate, rhythm, and stress, and older adults are less able to use envelope cues when noise-vocoding disrupts fine structure cues (Souza & Boike 2006; Sheldonet al. 2008; but also see Wingfield et al. 2000). More localized envelope cues contribute to the perception of phonemic contrasts based on the duration of speech segments, and older listeners need longer gaps and segment durations than younger listeners to perceive such contrasts (Gordon-Salant et al. 2006; Pichora-Fuller et al. 2006). Periodicity cues in the temporal fine structure based on the fundamental frequency and harmonic structure of speech are thought to contribute to voice quality and identity, clarity, as well as the ability to segregate concurrent voices, which is reduced in older listeners (Vongpaisal & Pichora-Fuller 2007). Converging evidence from psychoacoustic studies in older adults with normal pure-tone thresholds for their age (ISO 2000) and conducted using stimuli in the low-frequency range where their audiometric thresholds remain normal also point to age-related deficits in temporal processing for cues carried by the envelope as well as the fine structure (for a review see Pichora-Fuller & MacDonald 2008). In particular, reduced performance by older listeners on measures such as frequency difference limens (Abel et al. 1990), temporal fine structure sensitivity (Hopkins & Moore 2011), and specific patterns of binaural masking level differences (e.g., Pichora-Fuller & Schneider 1992) is consistent with the hypothesis that a loss of neural synchrony or reduced phase-locking may manifest as age-related declines in coding periodicity cues in temporal fine structure. Such age-related declines in auditory temporal processing seem to be consistent with physiological findings from studies using animal models (e.g., Khimich et al. 2005; Kujawa & Liberman 2009; Buran et al. 2010; Ison et al. 2010). Thus, further examination of this hypothesis in relation to speech understanding is warranted.

Pichora-Fuller et al. (2007) tested the extent to which jittering, or temporally distorting, the fine structure of the low-frequency components of speech, while controlling for the amount of spectral distortion, affected word-recognition performance in noise. Younger listeners with normal hearing (YN) were presented low-context sentences from the Revised speech perception in noise test (R-SPIN; Bilger et al. 1984) that were jittered in an effort to mimic the hypothesized effects of age-related neural dysynchrony. The results for younger listeners in the simulated auditory aging condition with low-context R-SPIN sentences were comparable with previous results for older adults with normal pure-tone thresholds through 3000 Hz to whom unaltered R-SPIN sentences had been presented (Pichora-Fuller et al. 1995). The findings supported their hypothesis that temporal jittering has a negative effect on speech performance in noise and that the temporal jittering algorithm could mimic the effects of temporal processing deficits associated with a loss of periodicity coding abilities in older adults with normal pure-tone thresholds through 3000 Hz.

In the present set of studies, two experiments were conducted to investigate the effects that two distortions (spectral smearing and temporal jittering) had on word recognition when the speech was presented in quiet or in multitalker babble over a range of SNRs. In quiet the distortions of the speech signal (i.e., temporal fine structure and possibly the envelope) are limited to the smearing or jittering processes that are applied, whereas babble introduces another type of distortion to the speech wave form that compounds to some degree the total distortion of the speech signal. The first experiment focused on evaluating word-recognition performance in YN using materials that were spectrally or temporally distorted. The second experiment was an extension of the first experiment; it aimed to evaluate the effects of age, hearing loss, and distortion type on word-recognition performance in quiet at two levels and in noise across a range of SNRs.

By comparing the results of the two experiments we could evaluate the extent to which the spectral and temporal distortions mimicked the effects of auditory aging. Our hypothesis was that the word-recognition performance of older listeners with normal hearing for pure tones (ONH) would be comparable with that of YN if audibility was unaltered but the temporal characteristics of the speech were degraded when words were presented to the younger listeners as had been found when SPIN-R materials were used (Pichora-Fuller et al. 2007). In addition, we hypothesized that to approximate the performance of older listeners with hearing loss (OHL) it would be necessary to distort speech spectrally when it was presented to either YN or ONH listeners. If the simulation were more effective when the distorted materials were presented to ONH listeners compared with YN listeners, then other age-related factors not captured by the implemented distortions would be implicated.

MATERIALS AND METHODS

Materials

In the following two experiments, the Northwestern University Auditory Test No. 6 (NU No. 6) materials were used to test word recognition in quiet (Tillman & Carhart 1966) and the Words-In-Noise (WIN) test materials were used to test word recognition in noise (Wilson 2003; Wilson et al. 2003; Wilson & McArdle 2007). Both the NU No. 6 and WIN materials were spoken by the same female speaker (Department of Veterans Affairs 2006). The original WIN test consisted of 70 words from the NU No. 6 lists that are presented in a six-talker babble at seven SNRs, ranging from 24 to 0 dB SNR in 4 dB decrements. The level of the babble was fixed. Subsequently, the 70-word version was divided into two complementary 35-word lists (List 1 and List 2) with five unique words presented at each of the seven SNRs (Wilson & Burks 2005).

Spectral and temporal distortions were applied to the NU No. 6 and WIN materials using the identical smearing and jittering algorithms described by Pichora-Fuller et al. (2007). The distortions (see in the subsequent section) were applied to the speech and babble materials after the SNRs were set and the materials had been mixed digitally. The materials were recorded on compact disc (CD; 44.1 kHz sampling rate, 16 bit; see Supplemental Digital Content 1-5, http://links.lww.com/EANDH/A75, http://links.lww.com/EANDH/A76, http://links.lww.com/EANDH/A77, http://links.lww.com/EANDH/A78, and http://links.lww.com/EANDH/A79, which show audio examples of the NU No. 6 word boat in the unaltered condition and in each of the distortion conditions.

Jitter • Jittering was applied only to a low-frequency band (0 to 1200 Hz). Briefly, a fast Fourier transform (FFT) was applied

to separate the stimuli into two bands, above and below 1200 Hz. Subsequently, an inverse FFT (IFFT) was used to convert the stimuli back to the time domain. The low-frequency band (0 to1200 Hz) was jittered (i.e., phase modulated using random time delays with an SD of 0.25 msec).* Last, the jittered low-frequency band and the unaltered high-frequency band were recombined.

Double-Jitter • In the double-jitter condition, the jitter algorithm was applied to the low-frequency band of the unaltered signal and was then applied again to the signal resulting from the first application of jittering.

Smear • Smearing was based on the simulation of the spectral distortion associated with cochlear hearing loss developed by Baer and Moore (1993, 1994).† Only the low-frequency portion of the stimuli (<1250 Hz) was smeared.

Jitter-Smear • For the jitter-smear condition, the jitter algorithm was applied to the unaltered stimuli. After this, the smear algorithm was applied to the jittered stimuli.

Participants

Different participants were recruited for the two experiments. Details about the participants will be described for each separate experiment. The YN were recruited locally in Upper East Tennessee. No audiology students served as participants. In the second experiment, a group of older listeners with normal or near-normal pure-tone thresholds also participated. These older listeners were recruited from the Mississauga area. The OHL in the second experiment were male Veterans recruited from the Audiology Clinic at the James H. Quillen Veterans Affairs (VA) Medical Center in Mountain Home, Tennessee. No participants had a history of middle ear or retrocochlear pathologies, and all were in good general health. All of the participants were native English speakers and were remunerated for their participation. These experiments were approved by the research oversight review boards at the respective facilities.

Common Procedures and Apparatus

At both the Mountain Home VA and University of Toronto laboratories, identical testing procedures and equipment were used. The participants completed pure-tone audiometry for octave frequencies of 250 to 8000 Hz and interoctave frequencies of 3000 and 6000 Hz (ANSI 2004). The speech materials were reproduced on a CD player (Sony, Model CDP-CE375), routed through an audiometer (Grason-Stadler, Model 61) to an earphone (TDH-50P) encased in a P/N 510C017-1 cushion. The nontest ear was covered with a dummy earphone. To avoid any ear effects, the stimuli were presented to the right ears of even-numbered listeners and the left ears of odd-numbered listeners. All testing was conducted in a double-wall sound booth.

EXPERIMENT 1

The purpose of this experiment was to determine the effects of jitter, double-jitter, smear, and jitter-smear distortions on word-recognition performances in quiet and in noise for YN. In this first experiment, the four distortions were applied to the speech and multitalker babble signals after the two signals had been mixed digitally.

METHODS

Participants

Sixteen younger naive listeners (mean age = 22.9 yr; SD = 2.7 yr) with normal hearing (≤20 dB HL) participated (ANSI 2004). The mean thresholds of the test ear were <10 dB HL from 250 to 8000 Hz.

Procedures

During a test session four randomizations of WIN List 1 and four randomizations of WIN List 2 were administered alternatively either in the order 1-2-1-2-1-2-1-2 (odd-numbered participants) or 2-1-2-1-2-1-2-1 (even-numbered participants). In the first half of the session, the following four distortion conditions were presented randomly: (1) jitter, (2) double-jitter, (3) smear, and (4) jitter-smear. In the second half of the session, an additional constraint was that the complementary WIN list was used for each of the four distortion conditions. In this manner, data from List 1 and List 2 were obtained for each of the four conditions with the possible effects of learning and fatigue randomized and spread across the four conditions. The babble used in the WIN test was presented at 80 dB SPL and the words were presented from 24 dB SNR (104 dB SPL) to 0 dB SNR (80 dB SPL) in 4 dB decrements. After the eight experimental conditions, the listeners were administered four 25-word NU No. 6 lists in quiet at 104 dB SPL, one in each of the four distortion conditions. The presentation order of the four NU No. 6 half-lists was random. After testing in quiet, the first WIN list (and distortion condition) given in the session was repeated so that learning and practice effects could be evaluated.

RESULTS AND DISCUSSION

Quiet

The first data columns in Table 1 list the mean percent-correct recognition performance (and SDs) obtained when the NU No. 6 words were presented in quiet for each of the four distortion conditions. Also shown in the table are data from a previous experiment on YN in the unaltered condition (Wilson et al. 2003). The performance on the unaltered WIN by the Wilson et al. participants was better

*Specifically, if $x(t)$ is the intact waveform, then the jittered waveform, $y(t)$, is a time-modulated version of $x(t)$ in which the time delay, δ, varies over time such that $y(t) = x[t-\delta(t)]$. There are two factors contributing to the degree of jitter. The first is the distributions of delays that might occur over a long period of time. The second is the rate at which the delay changes with time. The delay, $\delta(t)$, is determined based on the amplitude of a low-pass noise such that the SD of $\delta(t)$ is the root mean square (rms) amplitude of the low-pass noise. In the present study, an rms of 0.25 msec was used. Pichora-Fuller and Schneider (1992) found an internal jitter with an rms of 0.25 msec fit the data of older adults quite well. The rate at which the delay changes with time is determined by the bandwidth of the noise that is limited by the high-frequency cutoff such that a lower bandwidth results in a slower rate of change in $\delta(t)$.

†The stimuli were resampled to 16 kHz and divided into a series of frames, each frame consisting of 128 samples from the original stimuli and zero-padded with 64 zeros on each side of the frame to create a 256-point vector. A FFT was conducted on each frame and phase and power vectors were calculated independently. The low-frequency portion (<1250 Hz) of the power components of each were multiplied by a smearing matrix (procedure developed by Baer & Moore 1993). These smeared power components were then recombined with the unaltered phase components by an IFFT and the frames were recombined. In the present study, a smearing factor of 6 was used based on the results in a previous study by Pichora-Fuller et al. (2007) that matched spectral distortion of jittered and smeared stimuli from the R-SPIN test.

TABLE 1. The mean recognition performances (and 1 SD) of the 16 younger listeners with normal hearing from Experiment 1 in quiet (percent correct) and in noise (Spearman-Kärber 50% point) are listed for the four distortion conditions

	NU No. 6 Quiet		Spearman-Kärber		Polynomial	
Condition	%	SD	50% Point (dB SNR)	SD (dB)	50% Point (dB SNR)	Slope* (%/dB)
Unaltered†	96.8	3.3	4.1	1.4	2.9	6.5
Jitter	92.5	7.3	6.3	1.6	5.5	5.7
Smear	90.5	6.5	6.8	1.4	5.6	6.4
Double-jitter	88.8	5.1	6.9	1.7	5.8	5.6
Jitter-smear	87.3	7.8	8.2	1.7	7.0	5.2

The 50% points and slopes of the mean functions calculated from a 3rd-degree polynomial in Figure 1 also are listed.
**Slopes calculated at the 50% points on the polynomial functions in Figure 1.*
†From Wilson et al. (2003).

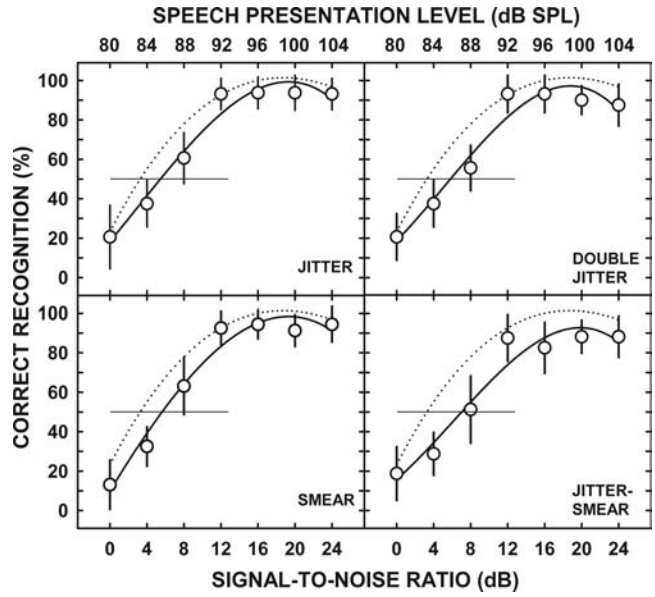

Fig. 1. Percent-correct recognition performance for words processed using various distortion algorithms shown as a function of signal to noise ratio (bottom abscissa) and speech presentation level (dB SPL, top abscissa) by the 16 younger listeners with normal hearing in Experiment 1. Error bars represent 1 SD. The solid line through the datum points represents the best-fit, 3rd-degree polynomial used to describe the mean data. The dotted line represents the mean performance for the unaltered Words-In-Noise test for the 24 younger listeners in the Wilson et al. (2003) study. The horizontal line crosses the 50% point on the functions.

than the performance by the participants in the distorted conditions in Experiment 1 of the present study. In the distorted conditions, the participants performed best in the jitter condition, followed by the smear condition, then the double-jitter condition, and they performed worst on the jitter-smear distortion condition. A within-subjects repeated-measures analysis of variance (ANOVA) was conducted to evaluate whether or not there was a significant effect of distortion type (jitter, double-jitter, jitter-smear, and smear) on performance for NU No. 6 words in quiet. The results revealed a significant main effect of distortion type, $F[3, 45] = 3.2$, $p < 0.05$. Post hoc evaluations using pairwise comparisons with Bonferroni adjustments for multiple comparisons were conducted ($p < 0.05$). Post hoc testing demonstrated that performance on the jitter-smear condition (87.3%) was significantly poorer than performance for the jitter condition (92.5%). The difference in performance between these two conditions was 5.2% or about a one-word difference on a 25-word list. No other significant differences were found for any other comparisons among the four distortion conditions in quiet.

Babble

Figure 1 illustrates the recognition performance (percent correct) on the WIN as a function of the SNR (dB) for each distortion condition. For reference purposes, each panel also contains the mean psychometric function (dotted line) obtained for the unaltered WIN from 24 YN pure-tone thresholds (Wilson et al. 2003). Note that the performance on the WIN test in all the distorted conditions is worse compared with the performance on the unaltered WIN test by listeners in the study of Wilson et al. (2003).

As shown in Table 1, the WIN data were evaluated in two ways, which yielded different information. Both the Spearman-Kärber equation and the polynomial fit approaches have been used in previous experiments with the WIN test and both are used in the present study so that our data can be compared with published data in prior studies conducted with these materials. First, the 50% points were calculated for each listener and each condition using the Spearman-Kärber equation (Finney 1952). This equation was used to describe the 50% point of the WIN functions for each individual enabling the calculation of group means and SDs as listed in the middle columns of Table 1. Second, the psychometric functions for the group means for each condition were fitted with a third-degree polynomial as illustrated with the solid lines through the data shown in Figure 1. The slopes at the 50% point of the mean functions depicted in Figure 1 were calculated using the first derivatives of the polynomials (see right columns of Table 1). The slopes were similar for jitter, double-jitter, and jitter-smear WIN conditions (~5.5%/dB), which were more gradual than the slopes for the smear and unaltered WIN conditions (~6.5%/dB).

As was found in quiet, recognition performance on the WIN was best for the jitter condition, followed by the smear, double-jitter, and jitter-smear conditions. An ANOVA using distortion (jitter, smear, double-jitter, and jitter-smear) as the within-subject factor indicated that there was a significant main effect of distortion for the 50% points determined by the Spearman-Kärber equation, $F[3, 45] = 19.7$, $p < 0.001$. Post hoc evaluation using Bonferroni corrections for multiple comparisons showed that only the performance on the jitter-smear condition (8.2 dB SNR) was significantly poorer than the performances on the other three conditions.

Additional analyses were conducted to determine whether or not the performances on the distorted WIN conditions were different from the performance on the unaltered WIN by YN from a previous study (Wilson et al. 2003). Separate t-tests (two-tailed) with p values adjusted for multiple comparisons (Bonferroni corrections) confirmed that performance for the unaltered WIN (Wilson et al. 2003) was significantly better than performance by the current listeners for WIN in jitter ($t = 4.5$, degrees of freedom [df] = 38, $p < 0.001$), smear ($t = 5.8$, df = 38, $p < 0.001$), double-jitter ($t = 5.7$, df = 38, $p < 0.001$), and jitter-smear conditions ($t = 8.2$, df = 38, $p < 0.001$).

Practice Effects

Recall that the first WIN condition presented in the session was repeated at the end of the session to evaluate possible practice effects. The initial mean 50% point was 7.1 dB SNR (SD = 1.6) and the final mean 50% point was 6.7 dB SNR (SD = 2.1). A paired *t*-test showed that this 0.4 dB improvement between the first and second administration of the identical WIN list/condition was not significant ($t = 0.6$, df = 15, $p > 0.05$).

Effect of Distorting the Speech or Babble

The similar performance on the WIN test when the stimuli were jittered and smeared was not consistent with the findings of a previous study using R-SPIN sentences in which YN performed significantly worse when the low-frequency components of speech were jittered than when the same components were smeared (Pichora-Fuller et al. 2007). In the previous study, however, only the speech signal was distorted and the babble was unaltered, whereas in Experiment 1 both the signal and the babble were distorted. To determine whether the discrepancies between the two studies might be caused by the distortion of the babble in the present study, a follow-up experiment was conducted. In this follow-up experiment, the effects of distorting neither the speech nor the babble, both the speech and the babble, or distorting one or the other were examined. In a different group of 16 younger listeners with average thresholds in the test ear <10 dB HL, four randomizations of WIN Lists 1 and 2 were administered in the following four conditions: (1) unaltered speech and unaltered babble, (2) unaltered speech with double-jittered babble, (3) double-jittered speech with unaltered babble, and (4) double-jittered speech with double-jittered babble. For the follow-up experiment, only one type of distortion was used; in Experiment 1 performance in the double-jitter condition was equivalent to performance with either the jitter or smear conditions, so of these three we selected the condition in which the distortion had been applied twice because it seemed more likely that the double-distortion condition might have a greater effect on performance than the single-distortion conditions. The distortions were applied to the speech and babble materials independently, after which the speech and babble were set to the appropriate SNRs and mixed digitally. The mean 50% points calculated by the Spearman-Kärber equation (±1 SD) for each condition in this follow-up experiment are listed in Table 2 along with the 50% points and slopes of the mean functions calculated from the third-degree polynomials. These results suggest that the primary degradations in word-recognition performance are a result of distorting the speech signal. There was a degradation in performance, although not significant, when the babble also was distorted. Thus, the difference found between the results of Experiment 1 and the Pichora-Fuller et al. (2007) study cannot be attributed to the babble in the WIN materials also being distorted in the present study.

EXPERIMENT 2

The purpose of this experiment was to evaluate word-recognition performances in ONH or near-normal pure-tone thresholds and in older listeners with pure-tone hearing loss, using the previously described temporally and spectrally distorted materials both in quiet and in babble. A secondary purpose was to assess the extent to which the recognition performances by the younger listeners in the distorted conditions evaluated in Experiment 1 resembled the performances with the unaltered materials for the two groups of older listeners. Furthermore, it was of interest to compare the performances on the distorted materials by the ONH or near-normal hearing to the performances on the unaltered materials by the OHL. The purpose of these comparisons was to assess the extent to which distorting the stimuli could mimic the effects of age-related differences in suprathreshold auditory processing or pure-tone hearing loss on word-recognition performance. In quiet, it was expected that audibility would be the primary factor affecting word recognition and that distortion would have relatively little effect on word recognition. In contrast, it was expected that distortion would have a greater effect on word recognition in noise because of poorer available fine structure cues, especially temporal fine structure cues, which would make segregating the speech from the competing babble more difficult for listeners.

METHODS

Participants

A total of 72 OHL (mean = 70.1 yr; SD = 7.5 yr) and 72 ONH (mean = 71.4 yr; SD = 4.4 yr) with near-normal pure-tone thresholds (defined as having 250 to 3000 Hz thresholds ≤25 dB HL) participated. All of the OHL group participants were males, whereas 56 females and 16 males were in the ONH group. Figure 2 shows the mean audiogram of the test ear for the two older listener groups and for the YN from Experiment 1.

TABLE 2. The mean 50% point (Spearman-Kärber) recognition performances in noise (and 1 SD) of the 16 younger listeners with normal hearing from the follow-up study to Experiment 1 are listed for the four conditions

WIN Condition		Spearman-Kärber		Polynomial	
Speech	Babble	50% Point (dB SNR)	SD (dB)	50% Point (dB SNR)	Slope* (%/dB)
Unaltered†	Unaltered†	4.1	1.4	2.9	6.5
Unaltered	Unaltered	4.6	1.5	3.6	6.3
Unaltered	Double-jittered	5.3	1.5	4.5	5.7
Double-jittered	Unaltered	6.6	1.5	4.9	5.9
Double-jittered	Double-jittered	7.0	1.4	5.7	5.9

The 50% points and slopes of the mean functions calculated from the 3rd-degree polynomial used to describe the data also are listed.
**Slopes calculated at the 50% points from the polynomial functions.*
†From Wilson et al. (2003).

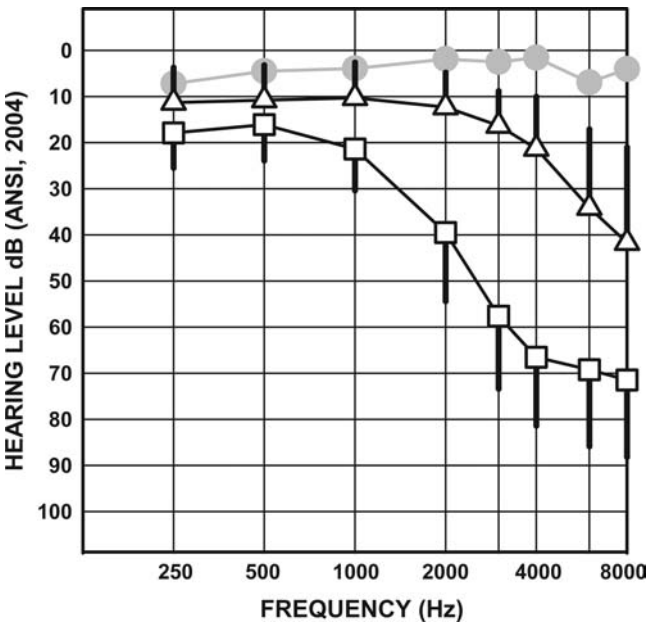

Fig. 2. The mean audiogram of the test ear in the younger listeners with normal hearing (circles) from Experiment 1 are illustrated along with the mean audiogram of the 72 older listeners with near-normal hearing (triangles) and 72 older listeners with hearing loss (squares) from Experiment 2. The error bars represent 1 SD.

On average, YN listeners had pure-tone thresholds <10 dB HL across all frequencies, the ONH listeners had normal pure-tone thresholds (i.e., ≤25 dB HL) through 4000 Hz and a mild hearing loss above 4000 Hz, whereas the older listeners with audiometric hearing loss had a moderately severe, high-frequency hearing loss above 1000 Hz.

Procedures

The participants were administered randomizations of Lists 1 and 2 of the WIN test in the following four conditions: (1) unaltered, (2) jittered, (3) smeared, and (4) combined jittered and smeared. As in Experiment 1, four randomizations of each of the two, 35-word WIN lists were used with no randomization repeated among the four listening conditions. The babble of the WIN was presented at 80 dB SPL, and the words were presented from 24 dB SNR (104 dB SPL) to 0 dB SNR (80 dB SPL). In addition to the WIN test, word-recognition performances in quiet were evaluated with NU No. 6 half-lists at the two levels, 80 and 104 dB SPL, which corresponded to the presentation levels of the words in the WIN paradigm at 0 and 24 dB SNR, respectively. To distribute equally any practice effects and any fatigue effects among the conditions, the presentation order of the 16 lists (eight 35-word WIN lists and eight 25-word NU No. 6 lists) was both counterbalanced and randomized among the participants.

RESULTS AND DISCUSSION

Quiet

As can be seen in Table 3, as expected, performance was best for the YN group (from Experiment 1) in all distortion conditions, followed by the ONH, and poorest for the OHL listeners. For both older listener groups, the pattern of performances across the four conditions was the same, with the best performance attained on the unaltered condition, followed by jitter, then smear, and last, the jitter-smear condition.

Within- and between-group differences in word-recognition performances in quiet among the conditions for the two older groups of listeners were evaluated using a mixed-model,

TABLE 3. The means (and SDs) obtained in the unaltered and distorted conditions by older adults with near-normal hearing and by older listeners with hearing loss on the NU No. 6 in quiet at two presentation levels are listed for four conditions in Experiment 2

	NU No. 6 Quiet (%)		SK 50%	50%	Slope
	80 dB SPL	104 dB SPL	dB SNR	dB SNR	%/dB
Unaltered					
Young normal*	—	—	4.6 (1.5)	3.6	6.3
Older normal	96.4 (5.0)	94.8 (5.7)	7.9 (2.1)	6.9	6.6
Older hearing loss	74.6 (19.0)	82.1 (13.4)	13.7 (3.5)	12.5	6.0
Jitter					
Young normal†	—	92.5 (7.3)	6.3 (1.6)	5.5	5.7
Older normal	88.2 (7.7)	83.0 (10.7)	10.8 (2.0)	9.5	6.4
Older hearing loss	48.6 (25.8)	57.7 (20.7)	16.9 (3.8)	15.7	4.1
Smear					
Young normal†	—	90.5 (6.5)	6.8 (1.4)	5.6	6.4
Older normal	83.9 (10.7)	73.6 (13.1)	11.2 (2.2)	9.8	6.8
Older hearing loss	41.5 (26.8)	51.9 (21.2)	17.3 (4.1)	15.9	4.2
Jitter-smear					
Young normal†	—	87.3 (7.8)	8.2 (1.7)	7.0	5.2
Older normal	77.3 (13.2)	60.9 (15.4)	13.6 (2.6)	11.7	5.8
Older hearing loss	30.5 (24.7)	32.6 (20.9)	19.5 (3.8)	—‡	—‡

Also listed are the mean 50% points obtained from the individual data with the Spearman-Kärber equation (SK 50%) and the 50% points and slopes at the 50% points calculated from the polynomial equations used to describe the mean data in Figure 3 for the older listeners in Experiment 2.
For comparison data from the young normals in Experiment 1 also are listed.
**From the follow-up study to Experiment 1.*
†From Experiment 1
‡Failed to reach 50%.

repeated-measures ANOVA.[‡] The between-subject factor was group (ONH and OHL). The two within-subjects factors were condition (unaltered, jitter, smear, and jitter-smear) and presentation level (80 and 104 dB SPL). The ANOVA revealed main effects for group, $F[1, 142] = 171.9$, $p < 0.001$, and for condition, $F[2.8, 396.3] = 551.1$, $p < 0.001$ (Greenhouse-Geisser correction), but not for level. The two-way interactions were significant (group by level, group by condition, and level by condition). The three-way interaction for level by group by condition also was significant, $F[2.9, 404.7] = 5.2$, $p = 0.002$ (Greenhouse-Geisser correction). Post hoc tests with Bonferroni corrections for multiple comparisons were used to evaluate differences in the three-way interaction.

As expected, for the OHL group, there was a significant difference in performance between the two presentation levels in quiet for all conditions (except for jitter-smear), with performance being better at the higher presentation level than at the lower presentation level. In contrast, the ONH group performed better when the words were presented at 80 dB SPL than when the words were presented at 104 dB SPL. The differences were observed for the four conditions, but were significant only for the smear and jitter-smear conditions. The higher presentation level may actually increase the audibility of the low-level components produced by the distortions for the ONH listeners, thus effectively changing the quiet condition into a signal in noise condition and accentuating the apparent effects of the distortions.

Babble

Figure 3 illustrates the psychometric functions for the two older groups tested in Experiment 2 and for the younger group tested in Experiment 1 in each of the four conditions. There are clear separations between the psychometric functions for the three listener groups, particularly in the dynamic portion of the functions. In addition, the asymptotes (i.e., plateau performance) are at a lower percent correct for the OHL in all conditions whereas they are at a high percent correct for the YN and the ONH listeners in the unaltered condition, with a reduction for the ONH listeners in the distortion conditions compared with the YN listeners.

Figure 4 shows mean performances (50% points Spearman-Kärber) for the three listener groups in the four conditions tested using the WIN materials. Figure 5 shows the same means as in Figure 4 as well as the individual 50% points for the four conditions plotted against each other. In the unaltered column (Figure 5, far left), almost all the datum points for all the listeners are above the line of equality, suggesting that almost all listeners performed better on the unaltered WIN than on any of the distorted conditions. Likewise, when comparing jitter-smear (bottom row) with all other conditions, almost all the datum points for all the listeners are above the line of equality, indicating that almost all listeners performed poorer on the jitter-smear condition than on all the other WIN conditions. When comparing the jitter and the smear conditions, all the datum points for all the listeners fall on or around the line of equality, suggesting nearly equal performance on these conditions for most individuals in each of the listener groups. Also shown in each panel of Figure 5 are Pearson r correlations among the conditions (top correlation). These correlations are all strong ($r > 0.88$), indicating that individual differences affect performance consistently across the conditions. Thus the pattern of results evident in the group means also is evident when looking at the patterns of individual data.

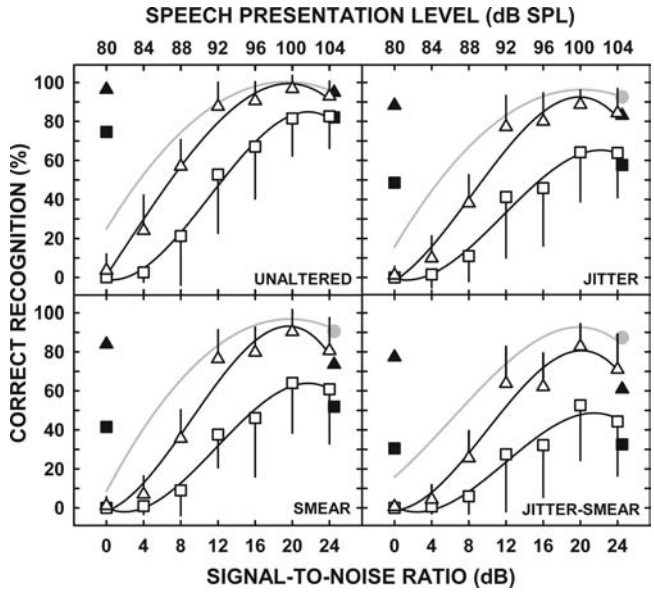

Fig. 3. Percent-correct recognition for words shown as a function of signal to noise ratio (bottom abscissa) and speech presentation level (dB SPL, top abscissa) for each condition (separate panels) are illustrated for the 72 older listeners with near-normal hearing (open triangles) and the 72 older listeners with hearing loss (open squares). The error bars represent 1 SD. The lines through the sets of datum points represent the 3rd-degree polynomial used to describe the mean data. For reference, the functions for younger listeners with normal hearing for the respective conditions from Experiment 1 are illustrated with a gray line (no symbols). The respective filled symbols represent mean recognition performance for word recognition in quiet in the given condition at the given speech presentation level (80 or 104 dB SPL) for each group.

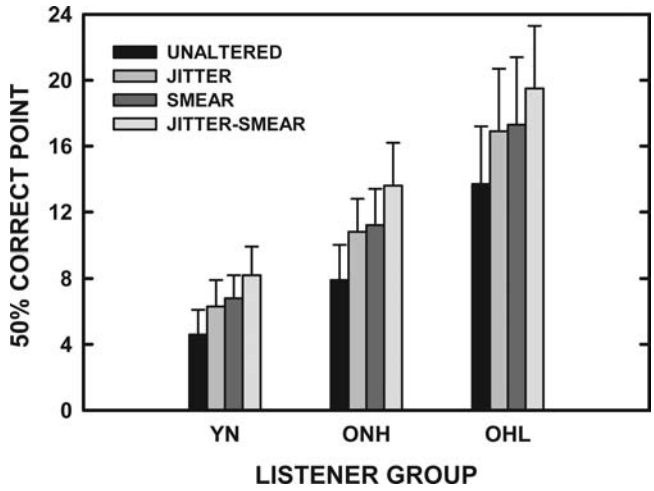

Fig. 4. The mean 50% points (dB SNR) calculated with the Spearman-Kärber equation for the unaltered (black), jitter (medium gray), smear (dark gray), and jitter-smear Words-in-Noise conditions (light gray) are illustrated as a function of the three listener groups. The error bars represent 1 SD.

[‡]The data also were analyzed using rationalized arcsine units rather than percent correct, but the results were the same; so the performances based on percent correct are reported here.

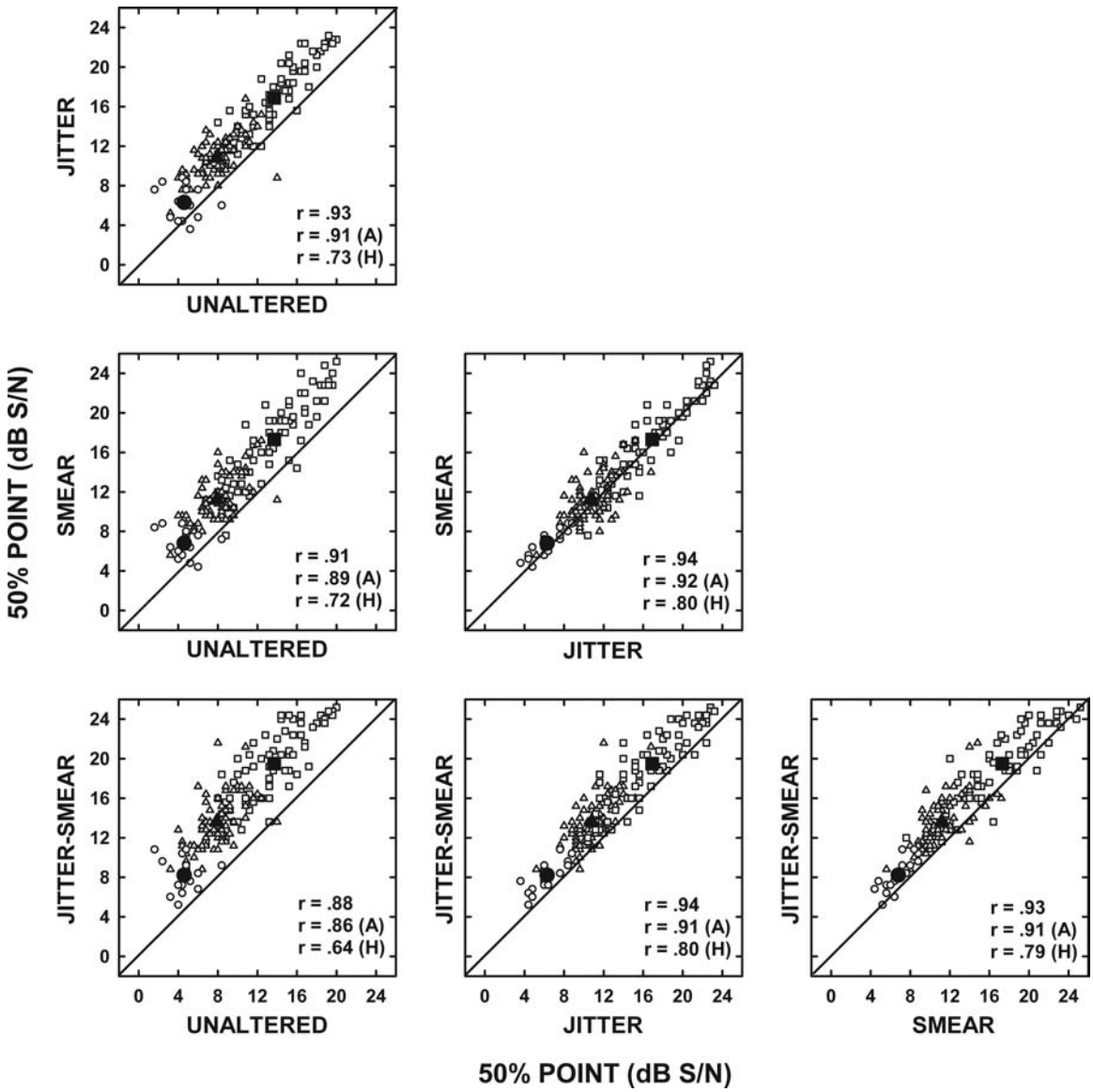

Fig. 5. The mean 50% points (dB SNR) calculated with the Spearman-Kärber equation for the individual listeners in each group are compared across conditions. In each panel, the individual datum points for the younger listeners with normal hearing from Experiment 1 (circles), older listeners with near-normal hearing (triangles), and older listeners with hearing loss (squares) are illustrated with open symbols. The large, filled symbols represent the group mean 50% points. The line of equality in each plot represents equal performance between the comparison conditions. In each panel, the bivariate correlations between the respective conditions are listed (top correlation). Also in each panel the partial correlations between the respective conditions when controlling for age (A) and high-frequency, pure-tone average (H) are shown.

The differences in mean recognition performances on the WIN test in the distortion conditions among the listener groups (listed in Table 3 and illustrated in Figures 3, 4, and 5) were evaluated with a mixed-model, repeated-measures ANOVA using the mean 50% points calculated with the Spearman-Kärber equation. In the analysis, the between-subject factor was group (YN, ONH, and OHL) and the within-subjects factor was condition (unaltered, jitter, smear, and jitter-smear). The results revealed a significant main effect of group, $F[2, 157] = 132.7$, $p < 0.001$, with all groups being significantly different from each other. There was a main effect of condition, $F[2.8, 432.4] = 244.0$, $p < 0.001$ (Greenhouse-Geisser correction), with performance in the unaltered condition being better than performances in the jitter and smear conditions, which were equivalent to each other, and performance in the jitter-smear condition being significantly worse than performances in all other conditions. Post hoc comparisons with Bonferroni corrections for multiple comparisons confirmed the aforementioned description of the differences among listener groups and conditions ($p < 0.05$). A significant group by condition interaction also was found, $F[5.5, 432.4] = 3.6$, $p < 0.005$ (Greenhouse-Geisser correction). As shown in Table 3, it is interesting to note that the performance of ONH group in the unaltered condition is similar to the performance of the YN group in the jitter-smear condition, and the performance for the OHL group in the unaltered condition is similar to the performance of the ONH group in the jitter-smear condition. This pattern of results related to the significant two-way interaction will be considered in the subsequent sections.

Additional analyses were conducted to tease apart the possible contributions that age and the audibility of high frequencies might have made to the results. In one analysis, partial correlations were conducted to investigate the influence that age and high-frequency, pure-tone average (HFPTA; average of the thresholds at 1, 2, and 4 kHz) have on the individual differences reflected in the strong correlations among WIN conditions shown in Figure 5. In the other analysis, bivariate correlations between WIN conditions and age and WIN conditions and HFPTA were calculated for each listener group. Last, an analysis of covariance (ANCOVA) was conducted to investigate if age or HFPTA were responsible for the effect of distortion on performance on the WIN test for the older listeners.

Two partial correlations were conducted between the distortion conditions shown in each panel of Figure 5, one controlling for age (A) and the other controlling for HFPTA (H). When controlling for age, the correlations between the conditions remained strong ($r = 0.86$ to 0.92) and similar to the bivariate correlations when no variable was controlled (top correlation in each panel). When controlling for HFPTA, the correlations were moderately strong ($r = 0.64$ to 0.80) but lower than the original bivariate correlations. These results suggest that age contributes less than HFPTA to the correlations among the conditions tested with the WIN materials. Even when HFPTA was controlled for, however, moderately strong correlations among the distortion conditions remained, suggesting that other factors besides HFPTA underlie the individual differences reflected in the correlations.

Two additional correlation analyses were conducted as a follow-up to the partial correlations stated earlier to evaluate further the relations between performance in the WIN conditions and age and HFPTA. In the first correlation analyses, for each listener group, we calculated the Pearson r bivariate correlations between age and performance in each WIN condition (see Figure 6). Overall, there was a significant and moderate correlation between age and all four WIN conditions (range: $r = 0.43$ to 0.53, $p < 0.001$). When evaluating the correlations separately for each listener group (as illustrated in the figure by the linear regression lines used to fit the data for each listener group), for the YN listeners there were no significant correlations with age; for the ONH group there were significant but weak correlations between age and performance in the distorted WIN conditions (range: $r = 0.26$ to 0.33, $p < 0.05$); and for the OHL group there were significant but weak correlations between age and all four WIN conditions (range: $r = 0.26$ to 0.38, $p < 0.05$). These results suggest that for both older groups age contributes to their performance in the WIN conditions when HFPTA is not controlled.

In the second follow-up correlation analyses, Pearson r bivariate correlations between HFPTA and performance in each WIN condition were calculated. For each listener group, Figure 7 illustrates the individual 50% correct points on the four WIN conditions as a function of HFPTA. Overall, there were significant and strong correlations between HFPTA and performance in each WIN condition (range: $r = 0.81$ to 0.86, $p < 0.001$). When evaluating the correlations separately for each listener group (as illustrated in the figure by the linear regression line used to fit the data for each listener group), for the YN listeners there were no significant correlations with HFPTA; for the ONH group there was a significant but weak correlation between HFPTA and performance in the jitter WIN condition ($r = 0.36$, $p = 0.05$); and for the OHL group there

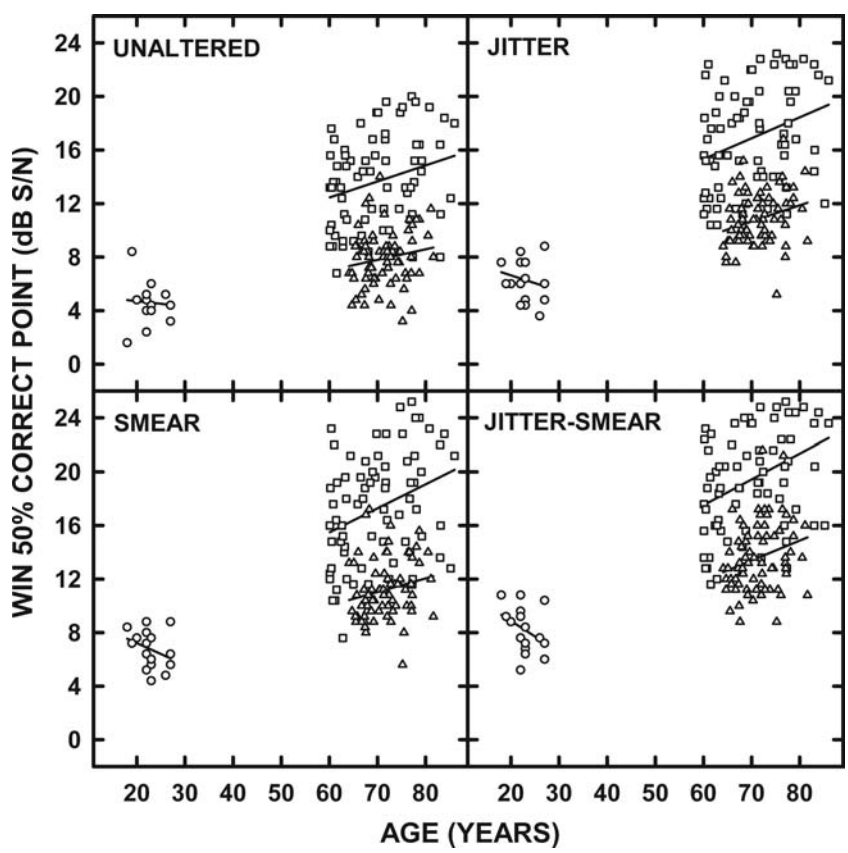

Fig. 6. The mean 50% correct points (in dB SNR) on the Words-in-Noise conditions are plotted separately in each panel as a function of age for each listener group (circles = YN, triangles = ONH, and squares = OHL). The linear regressions used to fit the datum points for each listener group also are illustrated.

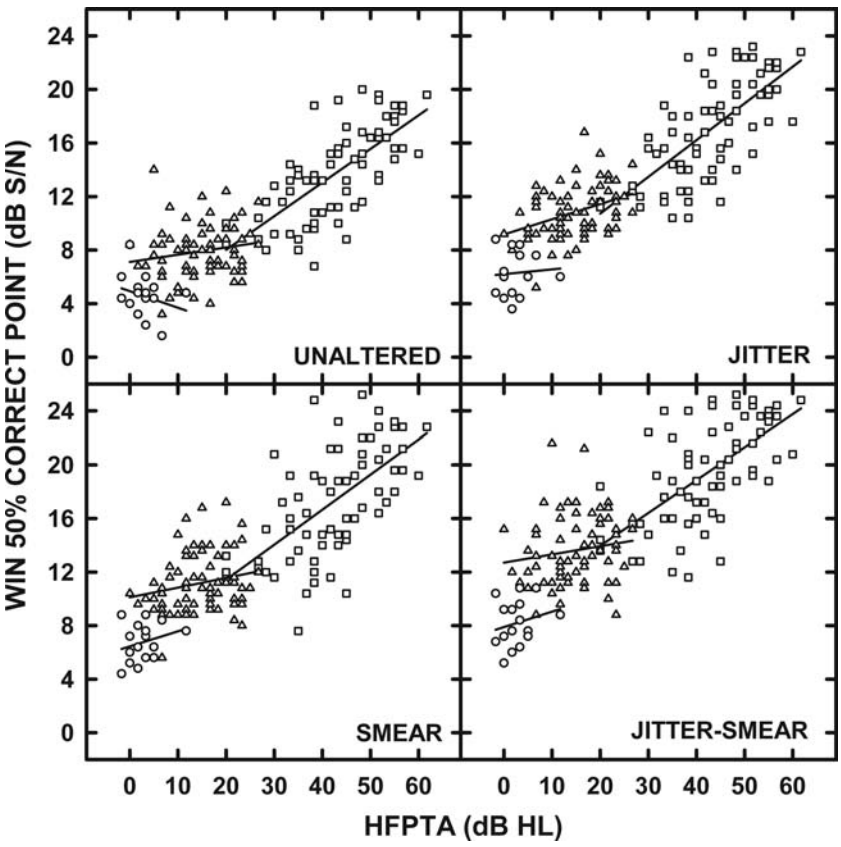

Fig. 7. The mean 50% correct points (in dB SNR) on the Words-in-Noise conditions are plotted separately in each panel as a function of high-frequency, pure-tone average for each listener group (circles = YN, triangles = ONH, and squares = OHL). The linear regressions used to fit the datum points for each listener group also are illustrated.

were significant and moderate correlations with HFPTA and all of the WIN conditions (range: $r = 0.63$ to 0.71, $p < 0.001$). Overall, these results suggest that HFPTA is a prominent factor contributing to the performance on the WIN for the OHL group, but not for the two normal-hearing listener groups, when age is not controlled.

In a final set of analyses, an ANCOVA was conducted to explore the contribution of age and high-frequency thresholds to the effect of distortion type on WIN 50% points. Because the assumptions of an ANCOVA are violated if the variables used as covariates also define the groups (Field 2009), the two older groups were combined into one group and the YN listeners were not included in the analysis. It is important to note that when the two older groups were combined there was no significant correlation between age and HFPTA ($r = 0.02$, $p > 0.05$). The significant main effect of distortion type found previously was found in a repeated-measures ANOVA conducted for the combined older group (F [2.8, 397.0) = 459.2, $p < 0.001$, Greenhouse-Geisser correction). Again, performance was best in the unaltered condition, followed by equivalent performances in the jitter and smear conditions, with the poorest performance in the jitter-smear condition. In the first ANCOVA, distortion type was the within-subjects variable and age was the covariate. The results revealed that there was no longer a main effect of distortion type. In the second ANCOVA, distortion type was the within-subjects variable and HFPTA was the covariate. The results revealed that the main effect of distortion type was preserved (F [2.8, 391.4] = 108.4, $p < 0.001$, Greenhouse-Geisser correction) as was the pattern of significant differences between the conditions. These ANCOVA results suggest that the effect of distortion type is eliminated when age is controlled, but the effect of distortion type remains when HFPTA is controlled. Thus, the effect of distortion type is tightly linked to one or more age-related factors but not HFPTA.

Distortion as a Simulation of Aging and Hearing Loss

On the basis of the results from the Pichora-Fuller et al. (2007) study, we expected that performance by younger listeners in the temporally distorted conditions would be similar to the performance by ONH listeners on unaltered WIN materials. Such similarities would suggest that the distortion altered the performance of the younger listeners in a way that resembled the effects of aging (rather than hearing loss because both YN and ONH groups had normal or near-normal audiometric thresholds; see Figure 2). We also expected that performance by the ONH listeners in the spectrally or temporally distorted conditions would mimic the performance of the OHL group on unaltered WIN materials. Such similarities would suggest that the distortions altered the performance of the ONH group in a way that resembled the effects of hearing loss (rather than age, which was the same for the two ONH and OHL groups).

Figure 8 is a compilation of various WIN psychometric functions that illustrate the comparisons of interest, including the performance of the younger listeners in distorted WIN conditions (star, diamond, and inverted triangle) and in the unaltered WIN condition of Experiment 1 (gray-filled circles). As described earlier, the distortions resulted in degraded performance for the younger listeners compared with their performance on unaltered WIN. When comparing the performances of the YN listeners in the distorted conditions in Experiment 1 with the performances of ONH listeners in Experiment 2 on the unaltered WIN (filled

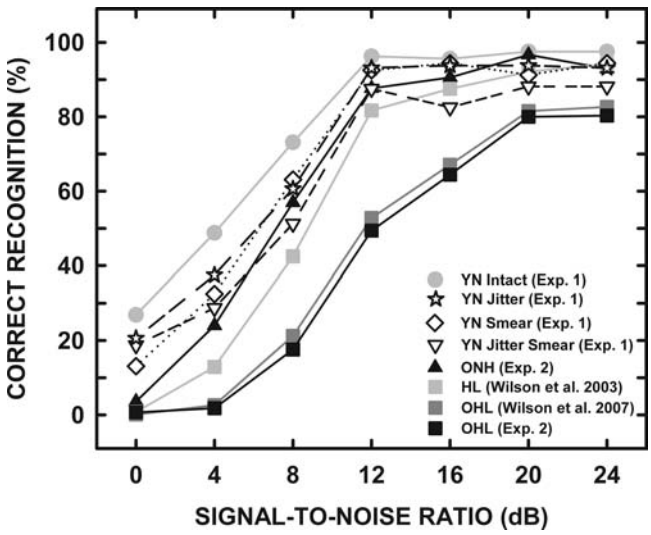

Fig. 8. The recognition performance (percent correct) of the 16 younger listeners with normal hearing in the jitter, smear, and jitter-smear Words-in-Noise (WIN) conditions from Experiment 1 are illustrated as a function of SNR ratio (dB) with open symbols and nonsolid lines. For reference in which to depict how well these distortions simulate performances of other listener groups for unaltered WIN, performances on unaltered WIN materials for the additional 16 younger listeners with normal hearing from Experiment 1 (light-gray circles), 24 middle-aged listeners with mild to moderate hearing loss from Wilson et al. 2003 (light-gray squares), and 72 older listeners with hearing loss from Wilson et al. 2007 (dark gray squares) also are illustrated. The mean performances on unaltered WIN for the older listeners with normal hearing (right-side up triangles) and with hearing loss (squares) from Experiment 2 are illustrated with filled black symbols.

black triangles), the performances on the distorted conditions mimic the performances of older listeners except at 0 dB SNR, where the performance of the older listeners in the unaltered condition falls below the performance of the younger listeners in the distorted conditions. Table 3 shows that the 50% points for the YN listeners in the jitter and smear conditions are better (by ~1.0 dB) than the 50% point for the ONH listeners in the unaltered WIN condition, whereas the performance by the YN listeners in the jitter-smear WIN condition is only 0.3 dB (7.9 versus 8.2 dB) worse than the performance of the ONH listeners in the unaltered WIN condition. Separate *t*-tests (using Bonferroni corrections for multiple comparisons) were conducted to evaluate whether or not there were differences in performance for the YN listeners in the jitter, smear, and jitter-smear WIN conditions compared with ONH listeners in the unaltered WIN condition. In terms of the 50% points, compared with the ONH group in the unaltered WIN condition, the performance of the YN listeners was significantly different for the WIN conditions with jitter ($t = -2.9$, df = 86, $p < 0.01$, two-tailed) and with smear ($t = -2.1$, df = 86, $p < 0.05$, two-tailed), but not with jitter-smear ($t = 0.5$, df = 86, $p > 0.05$, two-tailed). This finding confirms that the 50% point performance of the YN listeners in the jitter-smear WIN condition was comparable with the performance of the ONH listeners in the unaltered WIN condition, but YN listeners performed better in the jittered and smeared WIN conditions than did the ONH listeners in the unaltered WIN condition. Overall, when the WIN materials were both temporally and spectrally distorted, the performance of the YN listeners was reduced to the level observed for older listeners with clinically normal audiometric thresholds through 3000 Hz.

Figure 8 also shows how the results from the present study compare with the results obtained from two previous studies. The two lowest functions are from the 72 OHL tested in the unaltered WIN condition in Experiment 2 (black squares) and from 72 older listeners with moderate, high-frequency hearing loss tested in a prior study (dark gray squares; mean age = 70.8 yr; Wilson et al. 2007). These two groups of older listeners have similar pure-tone hearing losses and as seen in the figure they perform similarly on the WIN test but poorer than all other groups. The function in Figure 8 illustrated with light-gray, filled squares represents the mean WIN performances for 24 middle-aged listeners (mean age 58.5 yr) with mild, high-frequency hearing loss (Wilson et al. 2003). As seen in the figure, the performances of the YN listeners on WIN tests with combined jittering and smearing approximate the performances of the middle-aged listeners with hearing loss at high SNRs, but the performance of older listeners with moderate hearing loss is poorer.

The psychometric functions in Figure 9 illustrate the extent to which performances by ONH listeners in the distorted WIN conditions resemble the performance of the OHL listeners in the unaltered WIN condition. Because both groups are equivalent in age, the differences between them are assumed to be attributable to the effects of hearing loss arising from factors unrelated to aging per se. In Figure 9, the two shaded functions represent the performance of the ONH listeners (triangles) and the OHL listeners (squares) in unaltered WIN condition in Experiment 2. The functions with the open symbols represent performances on the three distorted WIN conditions by the ONH listeners. Recall that for the ONH listeners, the mean performances on the smear

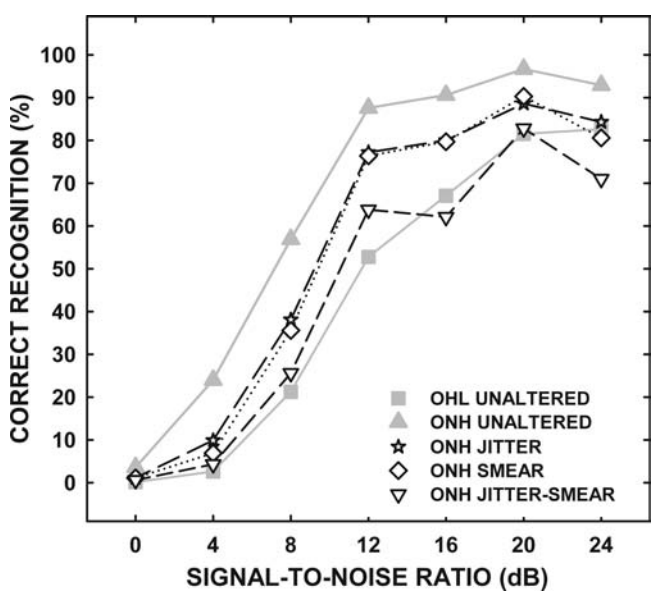

Fig. 9. The recognition performance (percent correct) of the 72 older listeners with near-normal hearing in the distorted Words-in-Noise (WIN) conditions are illustrated as a function of SNR ratio (dB) with open symbols and nonsolid lines. Performances on unaltered WIN materials for the 72 older listeners with near-normal hearing (light-gray triangles) and the 72 older listeners with hearing loss (light-gray squares) from Experiment 2 also are illustrated to depict how well these distortions simulate performances of older listeners.

and the jitter conditions were similar, with both being poorer than the performance on the unaltered condition and better than performance on the jitter-smear condition. The data in the figure indicate that jittering or smearing alone do not result in performance that approximates the performance of OHL listeners, but the jitter-smear combination does offer a good approximation. Separate *t*-tests (using Bonferroni corrections for multiple comparisons) were conducted to evaluate whether or not there were differences in performance between OHL listeners tested with unaltered WIN materials compared with ONH listeners tested in the jitter, smear, and jitter-smear conditions. In terms of 50% points, there were significant differences between the performance of the OHL group tested with unaltered materials compared with the performance of the ONH group tested in the jitter condition ($t = -6.0$, df = 142, $p < 0.001$, two-tailed) and in smear condition ($t = -5.1$, df = 142, $p < 0.001$, two-tailed), but not when the ONH group was tested in the jitter-smear condition (13.6 versus 13.7 dB). This pattern of group differences by condition suggests that overall, when the WIN materials were both temporally and spectrally distorted, the performance of ONH listeners approximated the effects on word recognition of hearing loss in their peers.

GENERAL DISCUSSION

In the present study, the effects of age and hearing loss on word recognition in quiet and in noise were evaluated by comparing the word-recognition performances of listener groups that differed in age or degree of hearing loss. The relative contributions of spectral and temporal fine structure cues were investigated by using test materials that were distorted. Speech was jittered in an attempt to simulate the effects of age-related changes in periodicity coding and it was spectrally smeared in an attempt to simulate the effects of broadened auditory filters. In this discussion the results of the present study are summarized and compared with the results of a previous study that used the same distortions in an attempt to simulate the effects of auditory aging on word recognition, followed by final comments on directions for future research concerning how different distortions of the fine structure of speech may affect speech intelligibility.

Effects of Distortions on Listener Groups Differing in Age or Hearing Loss

In the unaltered conditions, the YN and ONH listeners performed similarly in quiet, but not in noise, particularly at lower SNRs (see Figure 3), confirming the frequent report that age-related differences in speech-recognition performance are more pronounced in noise than in quiet (e.g., for reviews see CHABA 1988; Pichora-Fuller & Souza 2003). Furthermore, as expected, in both quiet and in noise, ONH listeners performed better in the unaltered conditions than the OHL listeners, presumably because the audibility of speech is reduced for the OHL listeners. The audibility of the unaltered WIN materials is illustrated in the top panel of Figure 10, which displays the spectra of the speech and babble signals along with the mean hearing thresholds of the three listener groups. Sufficient audibility was deemed to be at least 15 dB SL (e.g., Humes 2007). Using this criterion, for the OHL group, audibility was compromised for 4 kHz and higher when the presentation level was 104 dB SPL and for 1.6 kHz and higher when the presentation level was 80 dB SPL. For the ONH group, audibility was compromised only for 8 kHz when the presentation level was 104 dB SPL and for the 6.3 kHz and higher when the presentation level was 80 dB SPL.

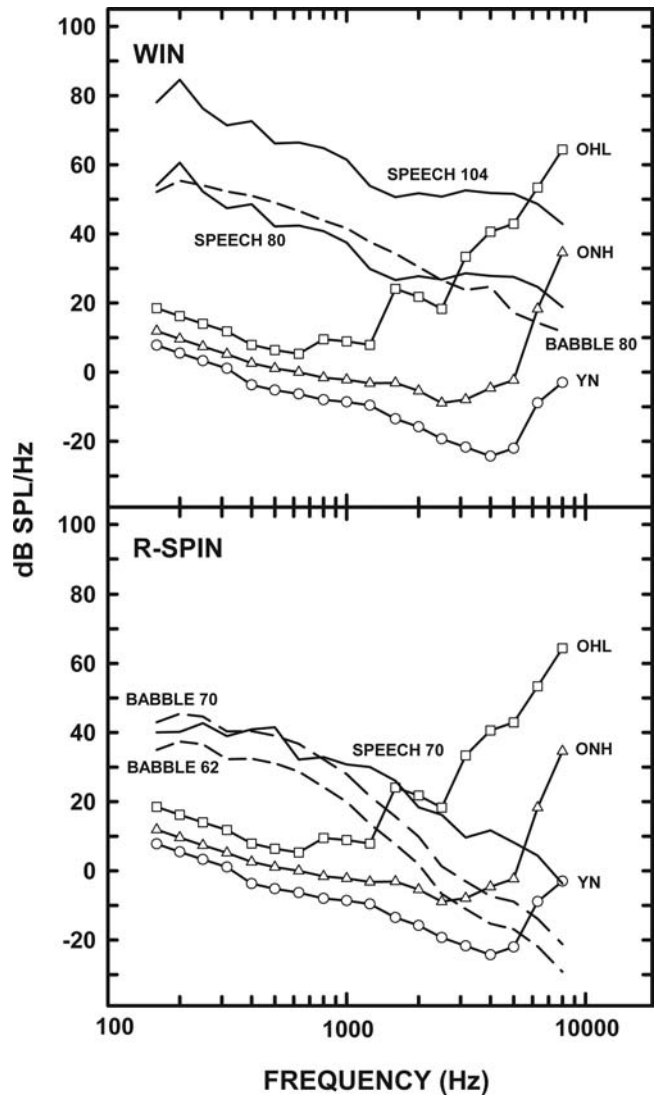

Fig. 10. The long-term average spectra for Words-in-Noise (WIN) speech and babble (top panel) and the revised speech perception in noise speech and babble (bottom panel) are plotted along with mean pure-tone thresholds for YN (circles), ONH (triangles), and OHL (squares) in both panels. The spectra were estimated using a 1/3 octave band procedure and then converted from band levels to spectrum levels. The pure-tone thresholds were measured in dB HL using TDH-50P headphones at standard audiometric frequencies and then converted to spectrum levels using Table 3 from ANSI (1997).

Both in quiet and in noise, performance by all listener groups decreased as the amount of distortion increased with the progression of the distortion conditions from single (jitter or smear) to combined (jitter-smear) distortion conditions. Furthermore, in noise, the results of the study suggest that OHL listeners are more susceptible to degradations in the speech signal than are ONH listeners, presumably because they are affected negatively by both the distortion and reduced audibility of the speech signal. Last, the ONH listeners are more susceptible to degradations in the speech signal in noise compared with YN listeners,

presumably because the ONH listeners are affected by distortion even though the effects of reduced audibility are minimal.

In the quiet distorted conditions, the performance of the OHL listeners improved as the presentation level of the speech increased, presumably because more speech information became audible as the level increased. It is interesting to note that word-recognition performances in the quiet distorted conditions for the ONH listeners were poorer when the materials were presented at 104 dB SPL than when the materials were presented at 80 dB SPL. There are several possible reasons that might explain why the ONH listeners performed poorer when the materials were presented at the higher presentation level compared with the lower presentation level. Poorer performance for the ONH listeners may have resulted from the upward spread of masking when the materials were presented at the higher presentation level (Studebaker et al. 1999). It is also possible that there may have been a compounding effect of the externally produced signal distortions and the internal distortions associated with auditory aging (Bocca & Calearo 1963; Plomp 1978). Compared with when speech is presented at lower levels, when speech is presented at high levels or in noise, temporal processing may be used by listeners with normal hearing to a greater extent whereas place-coding of frequency may be less useful (see Greenberg 1996); however, older listeners may not have access to this compensatory use of temporal processing (see Kujawa & Liberman 2009; Fitzgibbons & Gordon-Salant 2010). In addition, there may be an age-related increase in the internal distortions especially when the speech presentation level is higher. Such internal distortions would be consistent with the observation of rollover, or a decrease in word-recognition performance as the presentation level increases, which may occur when retrocochlear pathology is present as in some cases of presbycusis or when stimuli are distorted (Miranda & Pichora-Fuller 2002). Overall, the findings from the present set of experiments provide evidence that word-recognition abilities are influenced by external and internal factors such as presentation level, background noise, age, and hearing loss, but that the ways in which the external signal factors combine differs depending on the internal factors that are characteristic of the groups.

Effects of Distortion for WIN Versus R-SPIN Tests

In the present experiments with the WIN materials, the YN listeners performed similarly in the jittered and smeared conditions and performance by YN listeners approximated performance by ONH listeners on the unaltered WIN only when the speech materials were both jittered and smeared. These findings differ from the results of Pichora-Fuller et al. (2007) who found that temporally jittering the low-frequency components of the low-context R-SPIN test sentences reduced word-recognition performance in their YN listeners but spectral smearing of the same components did not. They also showed that the performance of YN listeners in the jitter condition was similar to that of ONH listeners tested earlier in another study. One difference between their study and the present study is that the materials were spoken by different talkers. The words used in the present study were recorded by a female talker (F_0 = 164 Hz), whereas the R-SPIN materials were recorded by a male talker (F_0 = 120 Hz). Insofar as jittering disrupts the periodicity of the signal, the consequences of jittering may depend on the fundamental frequency of the voice and the associated harmonic structure (Ives & Patterson 2008; Russo et al. 2011). The lower F_0 and more densely spaced harmonics of the male voice would normally offer richer periodicity cues to pitch than the female voice but they would be reduced to a greater extent by jittering. The speech spectra of the two materials also differed as shown in Figure 10. The R-SPIN speech materials contain more energy in the low-frequency region (<3000 Hz) compared with the WIN materials, whereas the WIN speech materials have more energy in the high-frequency region (>3000 Hz). Because the distortions were applied only to the low-frequency portion of the materials (<1200 Hz), the effects of the distortions may have been greater for the R-SPIN materials than for the WIN materials because a greater proportion of the R-SPIN materials would have been distorted.

Another difference between the materials used in the two studies is how well the amount of spectral distortion was matched when the materials were processed through the jitter and smear algorithms. It is inevitable that when the time wave form of a signal is altered, a byproduct of the change to the time wave form is spectral distortion. The jitter and smear algorithms both produce such spectral distortions in the materials; however, they do so differently. The amount of spectral distortion that is introduced by the two algorithms can be quantified and compared. In the study by Pichora-Fuller et al. (2007), by carefully matching the degree of spectral distortion introduced in the R-SPIN materials by both the jitter and smear algorithms, it was possible to demonstrate that the effect of temporal distortion introduced by jittering, rather than spectral distortion, was the source of the reduction in word-recognition performance. In the present experiments, the same jittering and smearing parameters were used as had been used in the Pichora-Fuller study without taking into account the possibility that differences in the speech stimuli might affect the amount of spectral distortion produced as a byproduct of the algorithms. Figure 11 illustrates as a function of frequency (0 to ~10,000 Hz) the amount of spectral distortion introduced by processing the NU No. 6 words in the WIN test using the jitter and smear algorithms. Note that the distortions do not alter the long-term average speech

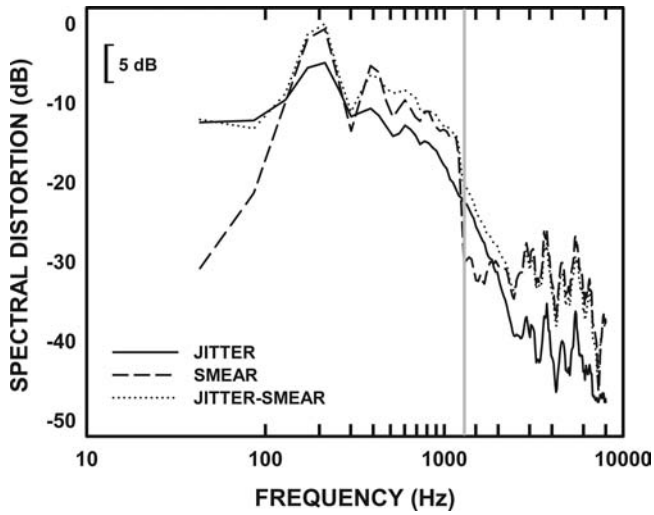

Fig. 11. The frequency distribution of spectral distortion introduced to the Words-in-Noise speech signal in the three processing conditions. The results are normalized relative to the peak distortion across all frequencies and conditions. The gray line represents the approximate cutoff frequency (1200 Hz) below which the distortions were applied.

spectrum. Further, the functions shown in Figure 11 do not illustrate the spectra of the processed signals, but rather the distribution of spectral distortion across frequency that is introduced in processing the signals using the algorithms.[§] As shown in the figure, there is greater spectral distortion produced in the smear and jitter-smear conditions compared with the amount of spectral distortion produced in the jitter condition. These differences in the amount of spectral distortion are consistent with the pattern of results observed across the distortion conditions. Thus, it seems likely that the apparent discrepancies between studies would not have arisen if the amount of spectral distortion had been equated between the jitter and smear conditions so that the unique contributions of temporal distortion in the WIN materials above and beyond spectral distortion could have been evaluated.

The effect of the spectral distortion resulting from processing the signals using the algorithms was examined further by calculating the SII for the four conditions (unaltered, jitter, smear, and jitter-smear). The SII predicts performance based on the spectra and presentation levels of the speech and noise and the pure-tone thresholds of the listener(s). The SII for each condition was calculated[¶] using the mean audiogram for each listener group (Figure 2).Within each of the three groups, the pattern of SII results across the various distortion conditions was the same. For each group, the differences between predicted intelligibility in the four processing conditions were on average less than 1%. Within each listener group, the similarity of the SII predictions across the four conditions is consistent with the claim that the distortions did not differentially affect the audibility of the signal across the spectrum. These results are also consistent with the view that the SII is not sensitive to the kinds of temporal and spectral distortions of the fine structure used in the present study (see Kates & Arehart 2005) and evidence suggesting that disruptions in the processing of fine structure cues can reduce the speech perception performance of older adults whether or not they have clinically significant audiometric hearing loss (Lorenzi et al. 2006; Pichora-Fuller et al. 2007; Lorenzi & Moore 2008; Moore 2008; Hopkins & Moore 2011).

In addition to the differences between the speech signals, it is also possible that differences between the prior and present study may have arisen because the babble in the R-SPIN and WIN materials differed both in terms of spectra and number of talkers, resulting in different masking effects. As shown in Figure 10, whereas the spectra of the babble used in the two tests are similar in the low-frequency region (<3000 Hz), in the high-frequency region (>3000 Hz) there is more energy in the WIN babble. Furthermore, a six-talker babble is used for the WIN test, whereas a 12-talker babble is used for the R-SPIN, and there may be slightly more informational masking relative to energetic masking when there are fewer competing talkers (e.g., Simpson & Cooke 2005). Nevertheless, these differences between the babbles in terms of the extent and type of masking they provided seem unlikely to explain why jittering had less effect on word recognition in the present study than in the prior study where masking was probably less challenging.

All of these factors may separately and in combination help to explain the differences between the performances of younger listeners in the two studies. The most likely explanation seems to be that differences in the speech of the two talkers interacted with the distortions.

Adequacy of Distortions to Mimic the Effects of Age and Hearing Loss on Word Recognition

The present experiments suggest that distorting the speech signal can result in reductions in word-recognition performance akin to those typically attributed to the effects of aging or moderately severe, high-frequency audiometric hearing loss. Follow-up analyses were conducted to examine the extent to which the strong correlations between performances in the various distortion conditions in noise were because of individual differences in age or HFPTA. These analyses suggest that both age and HFPTA were significantly correlated with performance in the different WIN conditions, but that these factors were not sufficient to provide a full explanation. Furthermore, follow-up analyses conducted using ANCOVA to examine the extent to which age and HFPTA contributed to the effect of distortion on performance in noise for the older listeners suggest that controlling for HFPTA did not alter the pattern of results whereas controlling for age eliminated the effect of distortion. Taken together, these analyses suggest that the differences between listeners evidenced in the strong correlations among the distortion conditions in noise were driven more by HFPTA and less by age and other factors, whereas the differences between distortion conditions evidenced in the within-subjects comparisons were driven more by age (or one or more age-related factors) than by HFPTA. The kinds of factors not related to age or HFPTA that might underlie the individual differences seen in the strong correlations between distortion conditions might be phonological processing ability or susceptibility to masking (i.e., aspects of suprathreshold processing that do not necessarily rely on fine structure cues). The kinds of factors related to age that might account for the differences between the distortion conditions seem most likely to involve the processing of the fine structure cues that were altered by the distortions. Thus, age-related differences in the use of fine structure cues

[§]The amount of spectral distortion was calculated using the same method as that used by MacDonald (2007), which compares the spectrotemporal energy distributions of the unaltered speech signal to the two distorted versions of the same signal. First the overall rms of both the jittered and smeared stimuli were normalized to equal that of the unaltered signal. Next, because the smearing algorithm introduces a small delay (4 msec), the smeared stimuli were temporally realigned with the unaltered signal. Using spectrograms plotted as images displaying the spectrotemporal energy distributions of the unaltered, jittered, and smeared signals, a pixel-to-pixel comparison was conducted. The magnitude of the pixel-to-pixel differences across time for the jitter compared with the unaltered conditions and for the smear compared with the unaltered conditions was averaged to yield the distribution of spectral distortion as a function of frequency. As an alternative to using spectrograms, Spectrotemporal Excitation Patterns (STEPs) estimate the spectrotemporal energy distribution based on properties of the human auditory system including models of the auditory filters and temporal windows. Calculations using STEPs yielded results that were similar to those obtained using spectrograms. Note that differences in spectral distortion calculated using this method capture time-specific differences in magnitude between the unaltered and distorted signals at each frequency, and such differences do not imply that there should be differences in the long-term average speech spectra if the negative and positive time-specific differences averaged over time cancel.

[¶]SII estimates were calculated using MatLab code written by Hannes Müsch (available from http://www.sii.to). A 1/3 octave band procedure was used with equivalent speech spectrum levels estimated from our speech signal (averaged over our 70 WIN words only; carrier phrase was removed for these calculations). Similarly, equivalent noise spectrum levels were estimated from our babble signal (averaged over the babble corresponding to the 70 WIN words only). The band importance function used was that for average speech.

could account for the differences between distortion conditions whereas HFPTA as well as aging and other aspects of processing not based on fine structure cues and not related to age could account for individual differences in performance that are stable across the distortion conditions in noise.

An examination of the effectiveness of the simulations in mimicking the effects of hearing loss and age provides another approach to evaluating their relative contributions to word recognition. For the WIN test, the performance of YN listeners in the jittered and smeared conditions mimicked the performance of ONH listeners in the unaltered condition at all but the worst SNRs, with performance at the 50% point being similar to that of the ONH listeners when YN listeners were tested in the jitter-smear condition. In the jitter-smear condition, the performance by the YN listeners did not approximate the performance by the OHL listeners in the unaltered conditions (see Figure 8); however, the performance by the ONH listeners in the jitter-smear condition did mimic the performance by the OHL listeners (see Figure 9). Although the YN and ONH groups both had normal or near-normal audiograms, it is interesting that the distortion was sufficient to render the performance of the older listeners similar to the performance of age-matched peers with hearing loss, but it was not sufficient for the younger listeners. The YN listeners may have been able to use the high-frequency speech components to a greater extent than the OHL listeners. Because the distortions were applied only to the low-frequency portion of the words (<1200 Hz), the specific deficits associated with high-frequency hearing loss were not addressed in the simulations. To simulate the effects of both aging and hearing loss in younger listeners, further compounding of the distortions tested in the present experiment could be evaluated. Alternatively, better simulations of the effects of hearing loss in older adults may be achieved by spectrally distorting or filtering the high-frequency components of the speech signal according to the degree of loss in the high-frequency range (see MacDonald et al. 2010 regarding the effect of temporally and spectrally distorting the high-frequency components of speech).

Previous studies have attempted to mimic aging or hearing loss through simulations with the presumption that external manipulations to the signal would adequately mimic the consequences of the internal processing associated with typical of cochlear hearing loss or the temporal processing declines believed to be associated with some types of auditory aging. Although we do not have a direct measure for determining whether or not the cochlear or neural processing of the distorted stimuli in YN listeners would be the same as the cochlear or neural processing of the unaltered stimuli in ONH listeners (i.e., simulation of auditory aging) or whether or not the cochlear or neural processing of the distorted stimuli in ONH listeners would be the same as the cochlear or neural processing of the unaltered stimuli in OHL listeners (i.e., simulation of cochlear hearing loss), nonetheless, such simulations may further our understanding of the underlying mechanisms that affect word-recognition performances in older adults with and without audiometric hearing loss. Overall, the results from the present experiments using simulations provide evidence that word-recognition abilities are influenced by intrinsic listener factors such as age and hearing loss, as well as by extrinsic stimulus factors such as the properties of the speech materials, presentation level, background noise, and other manipulations that alter the temporal fine structure of the speech signal.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the Department of Veterans Affairs (VA), Veterans Health Administration, and Office of Research and Development, Rehabilitation Research and Development (RR&D) Service. The work was supported by a Career Development Award (C6394W) to the first author, a Senior Research Career Scientist Award (C4256L) to the third author, and the Auditory and Vestibular Dysfunction Research Enhancement Award Program (C4339F) to the first and third authors, all sponsored by the VA RR&D, Washington, D.C. The portion of the research conducted in Toronto by the second author was supported by the Natural Sciences and Engineering Research Council of Canada [RGPIN 138472-05]. The authors are grateful to Ken Gabler, Kelly Koder-Carr, J. Kip Kelly, and Marianne Pelletier who assisted with the data collection. Huiwen Goy is acknowledged for her assistance with processing the WIN materials using the distortion algorithms.

This material is based on work supported in part by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, and Rehabilitation Research and Development (S.S. and R.H.W.) and the Natural Sciences and Engineering Research Council of Canada (M.K.P.).

Appropriate ethical and oversight committee approvals were obtained from both institutions before initiation of the study and all participants freely gave their informed consent before participation in this study. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article. The contents of this article do not represent the views of the Department of Veterans Affairs or of the U.S. Government.

Address for correspondence: Sherri L. Smith, James H. Quillen VAMC, Audiology (126), Mountain Home, TN 37684, USA. E-mail: sherri.smith@va.gov.

Received September 14, 2010; accepted November 8, 2011.

REFERENCES

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

American National Standards Institute (ANSI). (1997). *Methods for Calculation of the Speech Intelligibility Index*, S3.5. New York, NY: ANSI.

American National Standards Institute (ANSI) (2004). *Specifications for audiometers*, S3.6, New York, NY: ANSI.

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

Baer, T., & Moore, B. C. (1994). Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *J Acoust Soc Am*, *95*, 2277–2280.

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

Bocca, E., & Calearo, C. (1963). Central Hearing Processes. In J. Jerger (Ed). *Modern Hearing* 22 *Developments in Audiology* (pp. 337–370). New York: Academic Press.

Boothroyd, A., Mulhearn, B., Gong, J., et al. (1996). Effects of spectral smearing on phoneme and word recognition. *J Acoust Soc Am*, *100*, 1807–1818.

Buran, B. N., Strenzke, N., Neef, A., et al. (2010). Onset coding is degraded in auditory nerve fibers from mutant mice lacking synaptic ribbons. *J Neurosci*, *30*, 7587–7597.

Carhart, R. (1951). Basic principles of speech audiometry. *Acta Otolaryngol*, *40*, 62–71.

Ching, T. Y., Dillon, H., Byrne, D. (1998). Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification. *J Acoust Soc Am*, *103*, 1128–1140.

Ching, T. Y., Dillon, H., Katsch, R., et al. (2001). Maximizing effective audibility in hearing aid fitting. *Ear Hear*, *22*, 212–224.

CHABA (Committee on Hearing, Bioacoustics, and Biomechanics). (1988). Speech understanding and aging. *J Acoust Soc Am, 83*, 859–895.

Department of Veterans Affairs. (2006). Speech Recognition and Identification Materials, Disc 4.0 (Compact Disc). Mountain Home, TN: Department of Veterans Affairs.

Field, A. (2009). *Discovering Statistics Using SPSS* (*Introducing Statistical Method*). London: Sage Publications.

Finney, D. J. (1952). *Statistical Method in Biological Assay*. London: C. Griffen.

Fitzgibbons, P. J., Gordon-Salant, S. (2010). Behavioral Studies with Aging Humans: Hearing Sensitivity and Psychoacoustics. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, et al. (Eds). *The Aging Auditory System.* (pp. 111–134). New York: Springer.

French, N. R., & Steinberg, J. C. (1947). Factors governing the intelligibility of speech sounds. *J Acoust Soc Am, 19,* 90–119.

Frisina, D. R., & Frisina, R. D. (1997). Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. *Hear Res*, *106*, 95–104.

Gates, G. A., & Mills, J. H. (2005). Presbycusis. *Lancet*, *366*, 1111–1120.

Gordon-Salant, S., & Fitzgibbons, P. J. (1999). Profile of auditory temporal processing in older listeners. *J Speech Lang Hear Res*, *42*, 300–311.

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

Greenberg, S. (1996). Auditory Processing of Speech. In N. J. Lass (Ed). *Principles of Experimental Phonetics* (pp. 362–407). New York, NY: Mosby.

Hargus, S. E., & Gordon-Salant, S. (1995). Accuracy of speech intelligibility index predictions for noise-masked young listeners with normal hearing and for elderly listeners with hearing impairment. *J Speech Hear Res*, *38*, 234–243.

Hogan, C. A., & Turner, C. W. (1998). High-frequency audibility: Benefits for hearing-impaired listeners. *J Acoust Soc Am*, *104*, 432–441.

Hopkins, K., & Moore, B. C. (2011). The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. *J Acoust Soc Am*, *130*, 334–349.

Humes, L., & Dubno, J. (2010). Factors Affecting Speech Understanding in Older Adults. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, et al. (Eds). *The Aging Auditory System* (pp. 211–257). New York, NY: Springer.

Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *J Am Acad Audiol*, *18*, 590–603.

International Organization for Standardization (ISO). (2000). *ISO-7029: Acoustics Statistical Distribution of Hearing Thresholds as a Function of Age.* Switzerland: ISO.

Ison, J. R., Tremblay, K. L., Allen, P. D. (2010). Closing the Gap Between Neurobiology and Human Presbycusis: Behavioral and Evoked Potential Studies of Age-Related Hearing Loss in Animal Models and in Humans. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, et al. (Eds). *The Aging Auditory System* (pp. 75–110). New York, NY: Springer.

Ives, D. T., & Patterson, R. D. (2008). Pitch strength decreases as F0 and harmonic resolution increase in complex tones composed exclusively of high harmonics. *J Acoust Soc Am*, *123*, 2670–2679.

Kamm, C. A., Dirks, D. D., Bell, T. S. (1985). Speech recognition and the Articulation Index for normal and hearing-impaired listeners. *J Acoust Soc Am*, *77*, 281–288.

Kates, J. M., & Arehart, K. H. (2005). Coherence and the speech intelligibility index. *J Acoust Soc Am*, *117*(4 Pt 1), 2224–2237.

Khimich, D., Nouvian, R., Pujol, R., et al. (2005). Hair cell synaptic ribbons are essential for synchronous auditory signalling. *Nature*, *434*, 889–894.

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: Cochlear nerve degeneration after "temporary" noise-induced hearing loss. *J Neurosci*, *29*, 14077–14085.

Lorenzi, C., Gilbert, G., Carn, H., et al. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci USA*, *103*, 18866–18869.

Lorenzi, C., & Moore, B. C. J. (2008). Role of Temporal Envelope and Fine Structure Cues in Speech Perception: A Review. In T. Dau, J. M. Buchholz, J. M. Harte, et al. (Eds). *Auditory Signal Processing In Hearing Impaired Listeners* (pp. 263–272). Centertryk: 1st International Symposium on Auditory and Audiological Research (ISAAR 2007).

MacDonald, E. (2007). Temporal synchrony and aging. Unpublished Doctoral Dissertation, University of Toronto, Toronto, ON.

MacDonald, E. N., Pichora-Fuller, M. K., Schneider, B. A. (2010). Effects on speech intelligibility of temporal jittering and spectral smearing of the high-frequency components of speech. *Hear Res*, *261*, 63–66.

Mills, J. H., Schmiedt, R. A., Schulte, B. A., et al. (2006). Age related hearing loss: A loss of voltage, not hair cells. *Semin Hear, 27*, 228–236.

Miranda, T., & Pichora-Fuller, M. K. (2002). Temporally jittered NU-6 stimuli produce PI-PB rollover in young normal-hearing listeners. *J Am Acad Audiol, 13*, 50–58.

Moore, B. C. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol*, *9*, 399–406.

Moore, B. C., Glasberg, B. R., Simpson, A. (1992). Evaluation of a method of simulating reduced frequency selectivity. *J Acoust Soc Am*, *91*, 3402–3423.

Pichora-Fuller, M. K., & MacDonald, E. (2008). Auditory Temporal Processing Deficits in Older Listeners: From a Review to a Future View of Presbycusis. In T. Dau, J. M. Buchholz, J. M. Harte, et al. (Eds). *Auditory Signal Processing in Hearing Impaired Listeners* (pp. 291–300). Centertryk: 1st International Symposium on Auditory and Audiological Research (ISAAR 2007).

Pichora-Fuller, M. K., & Schneider, B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and old adults. *J Acoust Soc Am*, *91*(4 Pt 1), 2129–2135.

Pichora-Fuller, M. K., Schneider, B. A., 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., Schneider, B. A., Benson, N. J., et al. (2006). Effect of age on detection of gaps in speech and nonspeech markers varying in duration and spectral symmetry. *J Acoust Soc Am*, *119*, 1143–1155.

Pichora-Fuller, M. K., Schneider, B. A., Macdonald, E., et al. (2007). Temporal jitter disrupts speech intelligibility: A simulation of auditory aging. *Hear Res*, *223*, 114–121.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *Int J Audiol, 42*(*Suppl 2*), 11–16.

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J Acoust Soc Am*, *63*, 533–549.

Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *J Speech Hear Res*, *29*, 146–154.

Reed, C. M., Braida, L. D., Zurek, P. M. (2009). Review article: Review of the literature on temporal resolution in listeners with cochlear hearing impairment: A critical assessment of the role of suprathreshold deficits. *Trends Amplif*, *13*, 4–43.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond, B, Biol Sci*, *336*, 367–373.

Russo, F., Ives, T., Goy, H., et al. (2011). Age-related differences in melodic perception are probably mediated by temporal processing: Empirical and computational evidence. *Ear Hear*. Epublished ahead of print September 28, 2011.

Schmiedt, R. A. (2010). The Physiology of Cochlear Presbycusis. In S. Gordon-Salant, R. D. Frisina, A. N. Popper, et al. (Eds). *The Aging Auditory System.* (pp. 9–38). New York, NY: Springer.

Schneider, B. A., & Pichora-Fuller, M. K. (2001). Age-related changes in temporal processing: Implications for speech perception. *Semin Hear, 22*, 227–239.

Sheldon, S., Pichora-Fuller, M. K., Schneider, B. A. (2008). Effect of age, presentation method, and learning on identification of noise-vocoded words. *J Acoust Soc Am*, *123*, 476–488.

Simpson, S. A., & Cooke, M. (2005). Consonant identification in N-talker babble is a nonmonotonic function of N. *J Acoust Soc Am*, *118*, 2775–2778.

Souza, P. E., & Boike, K. T. (2006). Combining temporal-envelope cues across channels: Effects of age and hearing loss. *J Speech Lang Hear Res*, *49*, 138–149.

Studebaker, G. A., Sherbecoe, R. L., McDaniel, D. M., et al. (1999). Monosyllabic word recognition at higher-than-normal speech and noise levels. *J Acoust Soc Am*, *105*, 2431–2444.

Tillman, T. W., & Carhart, R. (1966). *An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words. Northwestern University Auditory Test No. 6*. Brooks Air Force Base, TX: USAF School of Aerospace Medicine Technical Report.

Vongpaisal, T., & Pichora-Fuller, M. K. (2007). Effect of age on F0 difference limen and concurrent vowel identification. *J Speech Lang Hear Res*, *50*, 1139–1156.

Willott, J.F. (1991). *Aging and the Auditory System: Anatomy, Physiology, and Psychophysics*. San Diego, CA: Singular Press.

Wilson, R. H. (2003). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *J Am Acad Audiol*, *14*, 453–470.

Wilson, R. H., Abrams, H. B., Pillion, A. L. (2003). A word-recognition task in multitalker babble using a descending presentation mode from 24 dB to 0 dB signal to babble. *J Rehabil Res Dev*, *40*, 321–327.

Wilson, R. H., & Burks, C. A. (2005). Use of 35 words for evaluation of hearing loss in signal-to-babble ratio: A clinic protocol. *J Rehabil Res Dev*, *42*, 839–852.

Wilson, R. H., & McArdle, R. (2007). Intra- and inter-session test, retest reliability of the Words-in-Noise (WIN) test. *J Am Acad Audiol*, *18*, 813–825.

Wilson, R. H., McArdle, R. A., Smith, S. L. (2007). An Evaluation of the BKB-SIN, HINT, QuickSIN, and WIN Materials on Listeners With Normal Hearing and Listeners With Hearing Loss. *J Speech Lang Hear Res*, *50*, 844–856.

Wingfield, A., Lindfield, K. C., 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.