

# Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners<sup>a)</sup>

Kara C. Schwartz,<sup>b)</sup> Monita Chatterjee, and Sandra Gordon-Salant

Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742

(Received 7 February 2008; revised 28 July 2008; accepted 10 September 2008)

The effects of spectral degradation on vowel and consonant recognition abilities were measured in young, middle-aged, and older normal-hearing (NH) listeners. Noise-band vocoding techniques were used to manipulate the number of spectral channels and frequency-to-place alignment, thereby simulating cochlear implant (CI) processing. A brief cognitive test battery was also administered. The performance of younger NH listeners exceeded that of the middle-aged and older listeners, when stimuli were severely distorted (spectrally shifted); the older listeners performed only slightly worse than the middle-aged listeners. Significant intragroup variability was present in the middle-aged and older groups. A hierarchical multiple-regression analysis including data from all three age groups suggested that age was the primary factor related to shifted vowel recognition performance, but verbal memory abilities also contributed significantly to performance. A second regression analysis (within the middle-aged and older groups alone) revealed that verbal memory and speed of processing abilities were better predictors of performance than age alone. The overall results from the current investigation suggested that both chronological age and cognitive capacities contributed to the ability to recognize spectrally degraded phonemes. Such findings have important implications for the counseling and rehabilitation of adult CI recipients.

© 2008 Acoustical Society of America. [DOI: 10.1121/1.2997434]

PACS number(s): 43.71.Lz, 43.66.Ts, 43.71.Es, 43.66.Sr [MSS]

Pages: 3972–3988

## I. INTRODUCTION

In recent years, cochlear implants (CIs) have proven to be a successful rehabilitative option for hearing-impaired (HI) listeners of all ages, including elderly individuals. In general, the surgical procedure is well tolerated among older individuals, and many patients report improvements in overall quality of life after implantation (Facer *et al.*, 1995; Vermeire *et al.*, 2005). Perhaps most importantly, several studies have demonstrated improvements in speech recognition in quiet when post-implantation scores are compared to pre-implantation scores (Kelsall *et al.*, 1995; Sterkers *et al.*, 2004).

Many authors report comparable performance across younger and older adult CI users (Kelsall *et al.*, 1995; Shin *et al.*, 2000; Pasanisi *et al.*, 2003); however, there is some evidence to suggest that age at implantation might influence overall speech recognition outcomes in this population (Dorman *et al.*, 1989, 1990; Gantz, 1993; Blamey *et al.*, 1996). Blamey *et al.* (1996) reported on a retrospective analysis and suggested that age at implantation significantly contributed to postimplantation speech recognition outcomes, particularly in those aged 60 and older. Other variables such as duration of deafness prior to implantation, device, hearing aid use, auditory-visual therapy, and preimplant auditory thresholds also contribute to speech perception abilities among adult CI listeners, thereby making contributions of

age at implantation alone less evident when evaluating performance among CI users (Parkin *et al.*, 1989; Dorman *et al.*, 1990; Blamey *et al.*, 1996; Rubinstein *et al.*, 1999; Leung *et al.*, 2005; Bodmer *et al.*, 2007; Green *et al.*, 2007).

### A. Causes of spectral distortion in cochlear-implant listeners and effects on speech perception

Regardless of the age of the listener, perception through a CI is greatly affected by a severe reduction in spectral cues. Several investigations have indicated that channel interaction, presumably attributed to electrical current spread, can greatly reduce the speech perception abilities of CI users and normal-hearing (NH) individuals listening to noise-band vocoded stimuli, regardless of the degree of spectral shifting (Shannon *et al.*, 1995; Dorman *et al.*, 1997; 1998; Fishman *et al.*, 1997; Friesen *et al.*, 2001). Based on these studies, it is estimated that CI recipients are listening through roughly six to eight spectral channels, at best, even though modern multi-channel CIs offer up to 22 intracochlear electrodes.

Further degradation is caused by a frequency-to-place mismatch or spectral shifting, as surgical limitations result in a shallow insertion depth of the electrode array. An electrode array is considered fully inserted when the most apical electrode approximates 25–30 mm from the base of the cochlea (Clark, 2003; Niparko, 2004). However, imaging studies have shown that a 20–25 mm insertion depth is typical, even among CI recipients who were reported to have a full insertion postsurgically (Ketten *et al.*, 1998; Skinner *et al.*, 2002). Assuming the average cochlear length of 35 mm, and applying it to Greenwood's (Greenwood, 1990) frequency-to-place equation, one could assume that the most apical electrode

<sup>a)</sup> Portions of this work were presented at the 2007 Aging and Communication Research Conference, held at Indiana University, Bloomington, Indiana.

<sup>b)</sup> Electronic mail: kschwartz@hesp.umd.edu

would be stimulating the frequency place corresponding to approximately 1000 Hz (for a 20 mm insertion) or 500 Hz (for a 25 mm insertion). If the lowest frequency of the speech signal presented to the electrode array was also 500 or 1000 Hz, then there would be no frequency-to-place mismatch. However, the frequency information presented to the apical most electrode (by the speech processor) contains frequency information as low as 100–200 Hz, therefore resulting in spectral shifting—commonly referred to as a “basal shift” of frequency information. The results of several investigations in both CI listeners and young NH individuals listening to CI simulations suggest that a frequency-to-place mismatch of 3 mm can cause significant decreases in speech perception abilities (Dorman *et al.*, 1997; Shannon *et al.*, 1998; Fu and Shannon, 1999). Although a reduction in spectral resolution can negatively affect the perception of various types of speech stimuli (i.e., phonemes, words, and sentences), vowel recognition appears to be very susceptible to spectral smearing and shifting (Dorman *et al.*, 1997; Friesen *et al.*, 2001; Baskent and Shannon, 2003) as the perception of vowels is highly dependent on the clarity of spectral cues (i.e., resolution of formant frequencies).

The effects of spectral degradation (number of channels and frequency-to-place mismatch) on phoneme recognition abilities have been thoroughly quantified among younger listeners, but not among middle-aged and older adults. It is known that learning and training factors result in significant improvements when listening with CI or noise-band vocoded stimuli (Rosen *et al.*, 1999; Svirsky *et al.*, 2004; Fu *et al.*, 2002, 2005), and there is some evidence to suggest that younger and older groups are able to adapt to spectrally distorted sentences over a short period of time (Peelle and Wingfield, 2005). However the immediate recognition of spectrally degraded phonemes by younger and older listeners has yet to be investigated.

## B. Potential interactions between aging and cochlear implantation

Even though CIs are undoubtedly beneficial to elderly HI listeners, there are reasons to suspect that the age of the adult listener might influence the ability to perceive spectrally degraded stimuli. First, CI processing results in severely diminished spectral cues, while temporal cues remain relatively unaltered. Therefore, CI users often rely on temporal cues to glean information about the target speech signal. This reliance on temporal cues seems to increase as spectral resolution decreases (Xu *et al.*, 2005; Nie *et al.*, 2006). Older listeners exhibit decreased sensitivity on several tasks measuring simple and complex temporal discrimination abilities, regardless of hearing status (Fitzgibbons and Gordon-Salant, 2001, 2004; Fitzgibbons *et al.*, 2007). Additional data imply that temporal processing deficits can arise as early as middle age (Snell and Frisina, 2000; Grose *et al.*, 2001; 2006; Lister *et al.*, 2002; Ostroff *et al.*, 2003). It is therefore hypothesized that such age-related changes in temporal processing abilities could affect speech perception when spectral cues are severely degraded, as when listening with a CI.

Second, a great deal of research has shown that older listeners show decreased speech recognition abilities when

the signal has been distorted. For example, older individuals, regardless of hearing status, show a decline in the ability to understand temporally compressed or reverberated speech, or speech that has been masked by modulated noise (Gordon-Salant and Fitzgibbons, 1993; 2004; Dubno *et al.*, 2002). A few investigations provide cursory evidence that sentence and phoneme recognition are influenced by the age of the listener when speech has been degraded in a manner similar to that produced by CI processing. Souza and Kitch (2001) measured sentence recognition abilities [Synthetic Sentence Identification (SSI) task] in young and old, NH and HI listeners, using undistorted stimuli and stimuli manipulated to contain a varying number of spectral channels (one or two) using the signal-correlated-noise (SCN) method (Schroeder, 1968). The SCN method mimics that of noise-band vocoding, in that it limits spectral cues while preserving temporal cues. The results indicated that the performance of the younger listeners exceeded that of the aged listeners by as much as 40 rationalized arcsine units (RAUs); the performance of the older participants listening to two channels of spectral information was worse than that of the younger participants listening to one channel of spectral information, regardless of hearing status. In a subsequent investigation, Souza and Boike (2006) measured 16-choice consonant recognition ability in both NH and HI listeners. Stimuli were processed to contain one to eight spectral channels using the SCN method (Schroeder, 1968). Performance within each channel condition was significantly related to the age of the listener: performance declined as the age of the listener increased.

Sheldon *et al.* (2008) measured the word recognition abilities of younger (aged 17–25 years) and older (aged 65–72 years) NH adults, by using noise-band vocoding methods and varying the number of spectral channels (1–16 channels). The results indicate that older adults needed approximately eight channels to achieve 50% accuracy, whereas younger adults only needed approximately six channels, when feedback was not provided and the number of channels was randomized across runs. Interestingly, significant age differences were absent if the presentation of stimuli across all listeners was not randomized, but rather the number of channels was increased with each run or if feedback was provided. These results underscore the role of training and feedback in the performance of older listeners. Peelle and Wingfield (2005) reported somewhat comparable performance between younger (18–21 years) and older (65–79 years) listeners when sentence recall was measured using 16-channel noise-band vocoded and spectrally compressed stimuli. Performance was tracked over 20 trials. Slight disparities were present during the initial trials with older listeners performing slightly but significantly worse than their younger counterparts; however, the average performance across the two age groups converged over the remainder of runs. Although these studies (Souza and Kitch, 2001; Souza and Boike, 2006; Peelle and Wingfield, 2005; Sheldon *et al.*, 2008) were not designed with the specific intent to investigate outcomes in cochlear implantation, results provide some evidence that age might influence the perception of spectrally degraded speech stimuli.

TABLE I. Average air-conduction audiometric thresholds (dB HL) within the three age groups. Standard deviations are indicated below the averages in parentheses.

Group	Ear	Frequency (Hz)									
		250	500	750	1000	1500	2000	3000	4000	6000	8000
YNH	R	7 (7.37)	7.5 (5.25)	7 (3.49)	9.5 (4.37)	9.5 (3.69)	6.5 (3.37)	11 (3.16)	4.5 (4.37)	6 (5.67)	6.5 (5.8)
	L	4 (4.59)	6.5 (4.74)	7.5 (4.29)	6.5 (4.11)	8 (4.21)	5 (3.33)	7.5 (5.89)	5.5 (3.68)	7 (4.21)	8 (4.83)
MNH	R	9.5 (4.97)	10.5 (8.64)	13 (5.86)	12.5 (6.34)	11 (6.14)	10 (6.23)	11.5 (4.74)	9.5 (4.97)	14.5 (6.85)	20 (8.18)
	L	8 (7.88)	11.5 (6.25)	10 (6.23)	12 (4.83)	10 (6.66)	9 (6.14)	13 (4.21)	11 (5.16)	13 (7.14)	20.5 (10.12)
ONH	R	8.5 (5.29)	9 (6.14)	13 (5.37)	14 (5.16)	11.5 (4.74)	10 (5.77)	16.5 (6.26)	10 (5.77)	16.5 (7.09)	16.5 (9.73)
	L	7 (6.32)	7.5 (6.34)	9.5 (3.68)	10 (4.08)	9.5 (4.97)	9.5 (4.37)	13.5 (5.8)	11.5 (6.25)	14 (5.16)	19.5 (7.19)

Lastly, cognitive status among adult CI users also appears to be an important factor in performance outcomes, because studies have found relationships between several cognitive measures and post-implantation speech recognition scores (Lyxell *et al.*, 1996; Heydebrand *et al.*, 2007). There is a tendency for some cognitive abilities (i.e., memory and speed of processing) to decline as the age of the adult increases; however, variability does exist within a given age (Verhaeghen and Salthouse, 1997). It is logical that a limited or degraded acoustic signal (bottom-up information) requires greater demand on cognitive (top-down) processing to accurately process and identify a speech signal. This relationship between bottom-up and top-down mechanisms is a popular model in both cognitive psychology and speech and hearing literature (see Wingfield and Tun, 2007 for review). It is possible that variation in senescent decline of cognitive ability could account for some of the disparities noted among the speech recognition performance of adult CI users.

### C. Summary and rationale

The effect of age on the immediate processing of spectrally degraded and shifted phonemes (e.g., the ability to correctly recognize and identify spectrally degraded stimuli without training or adaptation) has yet to be thoroughly investigated. It is reasonable to conjecture that age-related deficits in temporal processing might contribute to a decline in perception of spectrally degraded phonemes among older listeners, and recent evidence suggests that auditory temporal processing deficits begin as early as middle age (Snell and Frisina, 2000; Grose *et al.*, 2001; 2006; Lister *et al.*, 2002; Ostroff *et al.*, 2003). In the current study, the number of channels and degree of spectral shift were systematically varied to create different degrees of spectral distortion. The purpose of this study was to elucidate influences of age on the immediate perception of spectrally degraded vowel and consonant stimuli, by presenting stimuli to three different age groups: younger normal-hearing (YNH), middle-aged normal-hearing (MNH), and older normal-hearing (ONH) listeners. Because previous studies have shown an exceptional effect of spectral distortion on the recognition of

vowel stimuli, the current investigation places most of the focus on vowel perception, while superficially exploring the effects of aging on spectrally reduced consonants. The results are expected to shed light on difficulties experienced by adult recipients of CIs at initial activation of their device. Although it was not the focus of this investigation, a brief set of standardized cognitive tests was administered to account for possible influences of top-down processes.

## II. METHOD

### A. Participants

Thirty NH listeners participated in the current investigation. Participants were recruited based on their age at the time of testing; three groups of listeners were formed, with ten participants in each group: (1) YNH aged 21–25 years (mean age=22.4, SD=1.64), (2) MNH listeners aged 41–50 years (mean age=46, SD=2.94), and (3) ONH listeners (ONH) aged 61–71 years (mean age=65.2, SD=4.05).

For the purposes of this study, all individuals were required to have pure-tone thresholds  $\leq 20$  dB HL from 250–6000 Hz in one ear, whereas the other ear was required to have pure-tone thresholds  $\leq 25$  dB HL from 250–6000 Hz (ANSI, 2004). This modified criterion improved recruitment for the MNH and ONH groups. Audiometric data for all groups are provided in Table I. It is obvious that although the average pure-tone thresholds of the YNH group are slightly better than those to the MNH and ONH groups, all participants had *clinically* NH. Other audiometric criteria consisted of normal middle ear function, as measured by tympanometry (Roup *et al.*, 1998) and acoustic reflex thresholds (Gelfand *et al.*, 1990).<sup>1</sup>

All listeners reported good health and the absence of major physical and/or neurological impairments. Furthermore, all participants were screened for general cognitive awareness using the Mini-Mental State Examination (MMSE) developed by Folstein *et al.* (1975), which consists of 11 general tasks (30 points total). The authors note norms for the MMSE depend on the level of education and age of the individual. A t-score can be derived for each individual



based on these two variables; a t-score of less than 29 suggests cognitive impairment. All participants had t-scores well above this standard suggesting lack of significant cognitive impairment among the cohort.

## B. Stimuli

### 1. Phoneme identification

All speech stimuli were presented via the computer interface titled Computer-Assisted-Speech-Training (CAST©) Version 3.08 developed by Qian-Jie Fu of TigerSpeech Technology© at House Ear Institute (HEI). The tokens for the consonant and vowel recognition tests were digitized natural productions that were recorded at HEI for the CAST interface. The original recordings were preemphasized using a high-pass filter of 1200 Hz and were equalized in loudness. Vowel stimuli were originally recorded with a sampling rate of 16 000 Hz, while consonant stimuli were recorded using a 22 050 Hz sampling rate. There were 12 vowel stimuli consisting of the following 10 monophthongs presented in a /h/-vowel-/d/ context: heed: /h i d/, hod: /hɑ d/, head: h eɪd/, who'd: /h u d/, hid: /h ɪ d/, hood: /h ʊ d/, hud: /h əd/, had: /h æd/, heard: /hæd/, hawed: /hɔd/, as well as two diphthongs, hoed: /hɔud/ and hayed: /heid/. The consonant stimuli consisted of 16 phonemes /b, d, g, p, t, k, l, m, n, f, s, ʃ, v, z, ʒ, ð/, presented in an /aca/ context, where /c/=the consonant.

### 2. Noise-band vocoding

Noise-band vocoding was accomplished with TIGERCIS, also developed by Qian-Jie Fu (TigerSpeech Technology and HEI), using online signal processing within the CAST interface. For the purposes of this study, two variables were manipulated across experimental conditions: (1) the number of channels and (2) the lowest and highest frequencies of the analysis and carrier bands. The *analysis band* mimics the input audio range to the speech processor in a CI. The *carrier band* mimics the electrode array of the CI. The frequency content of the carrier band can be manipulated to simulate varying insertion depths according to the tonotopic organization of the cochlea. Manipulating the frequency information assigned to carrier and analysis band creates various degrees of spectral shifting (Fu and Shannon, 1999; Baskent and Shannon, 2003).

Stimuli were bandpass filtered into 4, 8, or 24 frequency bands using sixth-order Butterworth bandpass filters (24 dB/octave). The corner frequencies and bandwidths of the frequency bands were selected to simulate particular tonotopic locations and varied with the number of bands. Specific values were determined using the logarithmic equation provided by Greenwood (1990), assuming a 35 mm cochlear length.

The temporal envelope was extracted from each frequency band using half-wave rectification and low-pass filtering, with a cutoff frequency of 400 Hz (24 dB/octave) which is sufficient for good speech understanding (Shannon et al., 1995). The fine spectral content was replaced with a given number of noise carrier bands (4, 8, or 24), created from white noise by sixth-order Butterworth filters. The ex-

tracted temporal envelope was then used to modulate the corresponding spectral noise bands, resulting in decreased spectral information and preservation of temporal cues.

The frequency content of each stimulus was carefully selected to stay within the NH limits (upper limit of 6000 Hz) measured for each participant. The lowest frequency of all analysis bands (141 Hz, 31 mm from the base, approximately) was selected to approximate those commonly used in modern CI speech processors. Furthermore, the lowest frequency of the analysis band needed to be fairly low in frequency in order that the carrier band could be shifted, and still stay within the participants' NH range. The highest frequency used (5859 Hz, 9 mm from the base, approximately) was selected to be within the normal limits of hearing for all listeners. The division frequencies for the bandpass filters within each condition are listed in Table II according to the analysis or carrier band used in a particular condition. The division frequencies are shown for a 24 channel condition; however, division frequencies for the 8 and 4 channel conditions can be obtained from this table. Every third frequency in Table II corresponds to the division frequencies for the eight channel condition, and every sixth frequency corresponds to the division frequencies for the four channel condition. For example, for the "unshifted vowels and consonants," the division frequencies for an eight channel analysis band were 274, 468, 751, 1166, 1773, 2661, and 3960 Hz, and for a four channel analysis band were 468, 1166, and 2661 Hz. For all shifted conditions, the analysis band was fixed, while the carrier band was shifted to simulate varying insertion depths of the electrode array. The frequency band allocation was determined based on Greenwood's logarithmic function (Greenwood, 1990) for all experimental conditions.

### 3. Conditions

There was a total of 21 experimental conditions, which are described below and summarized in Table III. All test materials were stored on a computer disk and were output via custom software (TigerSpeech Technologies).

- (1) *Vowel recognition (12 choice)*. Vowel recognition was assessed using both unprocessed and noise-band vocoded stimuli. For the noise-band vocoded conditions, stimuli were manipulated to contain 4, 8, or 24 spectral bands (range of frequencies: 141–5859 Hz, corresponding to 31–9 mm from the base, respectively). Vowel recognition was also assessed when the analysis and carrier bands were mismatched to create spectral shifts of 3 or 5 mm. In shifted conditions, the analysis band was always fixed at 141–2684 Hz, while the carrier band was manipulated to create 3 mm (range of frequencies: 289–4409 Hz, 28 mm insertion depth) or 5 mm (range of center frequencies: 427–5859 Hz, 26 mm insertion depth) shifts. These two variables (number of channels and spectral shift) were varied systematically to create varying conditions of spectral degradation.
- (2) *Consonant recognition (16 choice, full set)*. Similar to vowel recognition, consonant recognition was measured using both unprocessed and noise-band vocoded stimuli

TABLE II. Division frequencies (Hz) for bandpass filters for all experimental conditions. Division frequencies are shown for 24 channel conditions; division frequencies for the 8 and 4 channel conditions can also be derived from the table.

Unshifted vowels and consonants																							
Corner frequency	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Analysis (31 mm from base)	180	214	274	330	395	468	550	644	751	872	1010	1166	1343	1545	1773	2032	2327	2661	3040	3471	3960	4514	5144
Carrier (0 mm shift=31 mm from base)	180	214	274	330	395	468	550	644	751	872	1010	1166	1343	1545	1773	2032	2327	2661	3040	3471	3960	4514	5144
Shifted vowels																							
Analysis (31 mm from base)	170	203	239	279	322	370	423	482	547	618	696	783	849	984	1101	1229	1370	1526	1698	1888	2098	2328	2583
Carrier (0 mm shift=31 mm from base)	170	203	239	279	322	370	423	482	547	618	696	783	849	984	1101	1229	1370	1526	1698	1888	2098	2328	2583
Carrier (3 mm shift=28 mm from base)	334	383	437	497	563	636	717	805	903	1011	1130	1261	1406	1566	1742	1936	2150	2386	2646	2933	3250	3599	3984
Carrier (5 mm shift=26 mm from base)	486	551	623	702	789	885	991	1108	1237	1379	1536	1709	1899	2110	2342	2598	2880	3191	3534	3913	4331	4791	5299
Shifted consonants (limited set)																							
Analysis (31 mm from base)	258	427	663	997	1469	2135	3078																
Carrier (2 mm shift=29 mm from base)	458	692	1018	1470	2097	2969	4179																

TABLE III. Parameters for all the experimental conditions.

	Stimuli	Number of channels	Analysis band (Hz)	Carrier band (Hz)	Shift (mm)
Unshifted	12 vowels	Unprocessed, 24, 8, 4	141–5859	141–5859	0
	16 consonants	Unprocessed, 24, 8, 4	141–5859	141–5859	0
Shifted	12 vowels	Unprocessed, 24, 8, 4	141–2864	141–2864	0
		24, 8, 4	141–2864	289–4409	3
		24, 8, 4	141–2864	427–5859	5
	6 consonants	Unprocessed, 8	141–4409	141–4409	0
		8	141–4409	233–5859	2

(4, 8, or 24 channels). The range of frequencies was 141–5859 Hz. No shifted conditions were presented using the full set of consonant stimuli. Instead, a limited set of consonants was used (see below).

- (3) *Consonant recognition (6 choice, limited set)*. Because many of the participants had a restricted range of sensitivity to the high frequencies, a limited set of stop consonants /b, d, g, p, t, k/ was used to measure the effect of a slight spectral shift. Furthermore, we thought it more efficient to limit the number of consonant conditions and place most of the focus of the current investigation on shifted vowel recognition. These consonants were specifically selected to limit consonants of other classes (fricatives and affricates) containing higher frequency information that, once shifted, would be mostly outside the range of NH for most participants. Although it is true that other classes of consonants (liquids, glides, and nasals) are also lower in frequency content, they exhibit vowel-like characteristics. Selecting these stop consonants enabled a slight (2 mm) shift of the carrier band. Performance was measured using unprocessed stimuli, eight channel noise-band vocoded stimuli (frequencies of analysis and carrier bands: 141–4409 Hz), and eight channel noise-band vocoded stimuli with a 2 mm shift (frequencies: analysis band=141–4409 Hz; carrier band: 233–5859 Hz).

#### 4. Cognitive measures

Verbal memory span, verbal working memory, and speed of processing abilities were measured using standardized tests available from the Wechsler Adult Intelligence Scale (WAIS-III) (Wechsler, 1997). Auditory versions of three memory subtests of the WAIS-III [forward digit span (FDS), backward digit span (BDS), and letter-number sequencing (LNS)] were recorded prior to administration.

The FDS test measures memory span; the listener is presented with a series of digits and asked to repeat the numbers back in the order they were presented. The first presentation consists of two digits (easy); however, the number of digits in each string increases with each presentation if the participant responds correctly. The BDS test is similar; however, the participant is required to repeat back the digits in the opposite order they were presented making it a test of both memory span and working memory abilities. The LNS test is a slightly different working memory task; a series of numbers and letters is presented to the listener in a random order.

The listener must first repeat back the digits in numerical order, and then the letters in alphabetical order. Similar to the FDS and BDS tests, the total number of letters and digits for the first presentation is small (three total), but increases if the participant continues to respond correctly. For all three tests, the individual must repeat the entire string in the correct order to receive credit; from these responses, a score is derived for each test. The total possible points for each test are as follows: FDS=16; BDS=14; LNS=21.

The speed of processing tasks included the symbol search (SEARCH) and digit symbol coding (CODE) tests. For the SEARCH task, two sets of symbols are shown on the test paper. The first set consists of two symbols and the second set consists of five symbols. Some symbols differ significantly from one another, whereas others are quite similar. The individual is required to examine the two symbols in the first column; if either of the two symbols matches one of the five symbols shown in the second column then the individual is to circle “Yes.” If neither of the two symbols appears in the five comparison symbols, then the individual is instructed to circle “No.” The second speed of processing task (CODE) consists of nine symbols, and each is associated with a number (1–9) shown in a key at the top of the page. Below the symbol key are several numbers listed in rows. The individual is required to write the corresponding symbol below each number by using the key at the top of the page. For both tasks, the listener is instructed to perform the test as quickly but as accurately as possible (2 minute time limit). There are 60 possible points for SEARCH and 133 possible points for CODE.

Stimuli for the verbal memory tasks were recorded with a Marantz Professional recorder (PMD660) using a 44 100 Hz sampling rate, then transferred and stored on a computer disk. The recorded materials were edited using ADOBE®AUDITION™ version 1.5. Each test was edited to contain 1 s interstimulus intervals (ISIs), thereby standardizing the timing of presentation across all participants. Edited stimuli were stored on computer disk, and then presented using the same equipment as described below under “Phoneme identification.” Speed of processing was measured using two subtests from the WAIS-III: digit symbol coding and the symbol search task. These written (visual) tests were administered in the standard fashion, requiring each participant to write his/her answers in the booklet provided.

## C. Procedure

### 1. Phoneme identification

All speech recognition testing was conducted in a double-walled sound treated booth. Stimuli were presented at 65 dBA in the free field through a single loudspeaker located at 0° azimuth. All speech stimuli were calibrated to a 65 dB SPL pure tone measured with a sound level meter, using a calibrated spot at the same location as the listener's ear. Calibration was performed through the CAST software, prior to each testing session. The CAST program presented the stimuli in a completely randomized order, within each experimental condition. The stimuli were output through a sound card, passed through a mixer (Rane SM 26B), and amplified (Crown D-75A), before they were presented through a single loudspeaker (Tannoy Reveal).

For vowel recognition measures, one repetition of each of the 12 vowels was produced by two speakers (one female and one male) for a total of 24 tokens (12 vowels  $\times$  2 talkers  $\times$  1 repetition). For consonant recognition measures, one repetition of each of the 16 consonants was produced by two speakers (one female and one male) for a total of 32 tokens per run (16 consonants  $\times$  2 talkers  $\times$  1 repetition). The limited set of stop consonants /b, d, g, p, t, k/ was drawn from the original set of 16 consonants. One repetition of each of the six consonants was produced by two speakers (one female and one male) for a total of 12 tokens (6 consonants  $\times$  2 talkers  $\times$  1 repetition) per run. No carrier phrase was used for either vowel or consonant stimuli.

Participants sat in a comfortable chair, facing a computer monitor. Each speech token was presented twice (1 s ISI), followed by 12 (vowel), 16 (consonant), or 6 (limited consonant) choices displayed on the computer screen in front of them, depending on the given condition. Each token was presented twice because (a) some of the conditions were extremely difficult to understand and (b) it was desirable to minimize contributions of inattentiveness as much as possible. After hearing both stimulus presentations, listeners were asked to select which speech token they thought they heard by selecting one of the choices displayed on the computer monitor in front of them (clicking on the icon using a computer mouse).

For the vowel stimuli, the buttons appeared as follows: HEED, HOD, HAWED, HOED, WHO'D, HAYED, HEARD, HUD, HID, HEAD, HOOD, and HAD. For the consonant stimuli, the buttons appeared as ACA, where "C" was the consonant for each stimulus. The choices appeared in very large black font within large gray squares and always appeared in the same order. All listeners reported that they were able to clearly see each of the choices displayed on the screen. Once the participant completed a selection, the next stimulus was presented through the loudspeaker with a delay of 1.5 s.

The conditions were presented in a randomized order, but each condition was presented five times (block randomized by runs). The task was self-paced, as listeners were not required to respond within a certain period of time. No feed-

back was provided. Testing was completed in approximately 3–4, 2–2.5 h sessions, with frequent breaks provided at the listener's discretion.

All participants were familiarized with the software, interface, and procedure using unprocessed stimuli prior to administration of the experimental conditions. Each listener performed practice runs with the unprocessed stimuli and heard a sample of a noise-band vocoded phoneme (eight channel, no shifting) before beginning the actual trials; each listener was required to obtain at least a 90% correct performance score on vowel and consonant recognition tasks with unprocessed stimuli prior to continuing in the study.

### 2. Cognitive measures

Cognitive subtests were always administered following the completion of all phoneme recognition testing, and were either conducted on a separate day or following a 15–20 min break after the completion of the phoneme testing. Similar to the procedure described above, participants were tested in a double-walled sound treated booth and sat in a comfortable chair. Stimuli for the FDS, BDS, and LNS tests were presented through a single loudspeaker and listeners responded using a microphone placed in the sound booth; responses were recorded by the test administrator who sat outside the sound booth. For the speed of processing tasks, listeners were required to write their answers in the test booklet provided. Order of presentation was the same across all listeners (SEARCH, CODE, FDS, BDS, LNS).

## III. RESULTS

### A. Phoneme identification

Prior to statistical analyses, percent correct scores for phoneme recognition were converted to RAUs to standardize variances across the mean (Studebaker, 1985). The resulting RAU ranges were  $-0.32$  to  $113.64$  for vowels,  $-3.31$  to  $114.87$  for consonants (full set), and  $8.79$ – $109.94$  for consonants (limited set). Performance across all five runs was averaged to obtain a mean RAU within each condition for each participant.

#### 1. Analysis and carrier bands matched (unshifted)

Figures 1 and 2 show the vowel and consonant recognition scores (in RAUs), respectively, across the three groups as a function of number of channels (abscissa). Similar to previous studies (Dorman *et al.*, 1997; Fu *et al.*, 1998), it is apparent that compared to consonant recognition, vowel recognition was more susceptible to spectral degradation, regardless of the age of the listener. Furthermore, performance is slightly dependent on the age of the listener, as phoneme recognition ability tended to decrease as the age of the listener increased. A mixed analysis of variance (ANOVA) with one within-subject factor [number of channels (four levels: unprocessed, 24, 8, 4)] and one between-subject factor (group) was used to analyze both vowel and consonant recognition data. The results for the vowel recognition tasks indicate a significant main effect of number of channels [ $F(3, 81) = 723.8$ ,  $p < 0.001$ ], and a significant main effect of



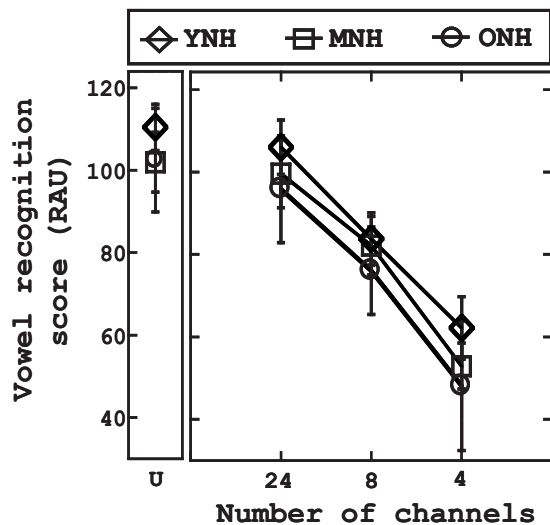


FIG. 1. Vowel recognition (unshifted) as a function of the number of channels for each group; "U" represents data obtained when stimuli were unprocessed. Recognition scores (RAUs) are plotted along the ordinate. Chance performance was 8.3% (6.2 RAUs). Error bars represent  $\pm 1$  standard deviation from the mean score within each group.

group [ $F(2, 27) = 5.47, p < 0.05$ ]. Multiple pairwise comparisons (dependent t-tests with Bonferroni adjustment) were statistically significant ( $p < 0.001$ ), suggesting that performance decreased with each decrease in the number of channels (full > 24 > 8 > 4). *Post hoc* testing (Tukey HSD) suggests that the performance of the YNH was significantly better than that of the ONH group ( $p < 0.01$ ); no other comparisons were significant.

Analysis of the consonant recognition data revealed that the assumption of sphericity was violated, so the Greenhouse-Geisser correction was used to interpret results. The results indicate a significant main effect of channels [ $F(2.143, 57.86) = 751.69, p < 0.001$ ] and a significant main effect of group [ $F(2, 27) = 10.67, p < 0.001$ ]. Similar to the

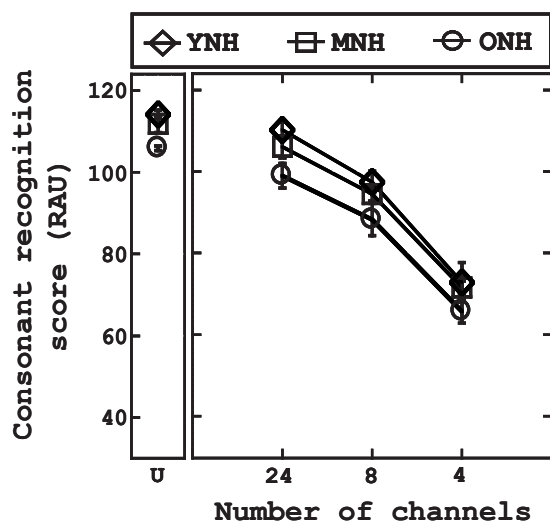


FIG. 2. Consonant recognition (unshifted) as a function of the number of channels for each group; "U" represents data obtained when stimuli were unprocessed. Recognition scores (RAUs) are plotted along the ordinate. Chance performance was 6.2% (2.75 RAUs). Error bars represent  $\pm 1$  standard deviation from the mean score within each group.

results reported for unshifted vowel recognition, pairwise comparisons using Bonferroni adjustments (dependent t-tests) suggest that consonant recognition performance decreased significantly with each decrease in the number of channels ( $p < 0.001$ ). *Post hoc* testing (Tukey HSD) revealed that the performance of the ONH group was significantly worse than that of the YNH and MNH groups ( $p < 0.05$ ) at each level of the independent variable (number of channels).

## 2. Vowel identification with spectral shifting

The results for shifted vowel recognition are shown in Fig. 3. Each panel represents results from each of the age groups. The abscissa indicates the degree of spectral shift (in mm), while the ordinate is the average RAU score. Results with unprocessed stimuli are also shown for reference. Overall, it is apparent that spectral shifting results in further declines in vowel recognition performance across all groups, compared to unshifted performance, which is consistent with previous reports (Fu and Shannon, 1999). Although slight age differences were noted when stimuli were unshifted, the addition of a spectral shift seems to enhance disparities in performance across the age groups, with YNH listeners' performance being far more robust to shifting when compared to the MNH and ONH groups. It is also noted that performance on the unshifted conditions depicted in Fig. 3 is slightly worse than those unshifted conditions shown in Fig. 1. This is likely due to the limited bandwidths used for the shifted vowel conditions (refer to Tables II and III for specific bandwidths used in both cases).

The shifted vowel recognition data were analyzed using a mixed ANOVA with two within-subject factors [number of channels (three levels: 24, 8, or 4) and degree of spectral shift (three levels: 0, 3, or 5 mm)] and one between-subject factor (group). The results suggest a significant three-way interaction between number of channels, degree of shift, and group ( $p < 0.05$ ). Follow-up analyses (using pairwise comparisons with Bonferroni adjustments and Tukey HSD, as appropriate) revealed complex interactions between phoneme recognition, spectral shifting, and the age of the listener. These differences are briefly highlighted below.

Pairwise comparisons revealed that the interaction between shift and number of channels depended on the age of the listener; more specifically, each decrease in number of channels caused a significant decline in performance across all three age groups for the 0 and 3 mm shifted conditions ( $p < 0.0167$ ). However, when a severe spectral shift was present (5 mm), performance depended on the age of the listener: the performance of the YNH group significantly improved as the number of channels increased from 4 to 8 and from 4 to 24 channels. Performance among the MNH and ONH groups, however, remained stable regardless of the number of channels in the 5 mm shifted condition. Further analysis revealed that the YNH listeners performed significantly better than the MNH or ONH groups in all of the 3 and 5 mm shifted conditions ( $p < 0.05$ ). However, for the unshifted conditions, a different pattern emerged with all three groups performing significantly differently than one another in the four channel condition (YNH performing the best and ONH performing the worst). Furthermore, the de-



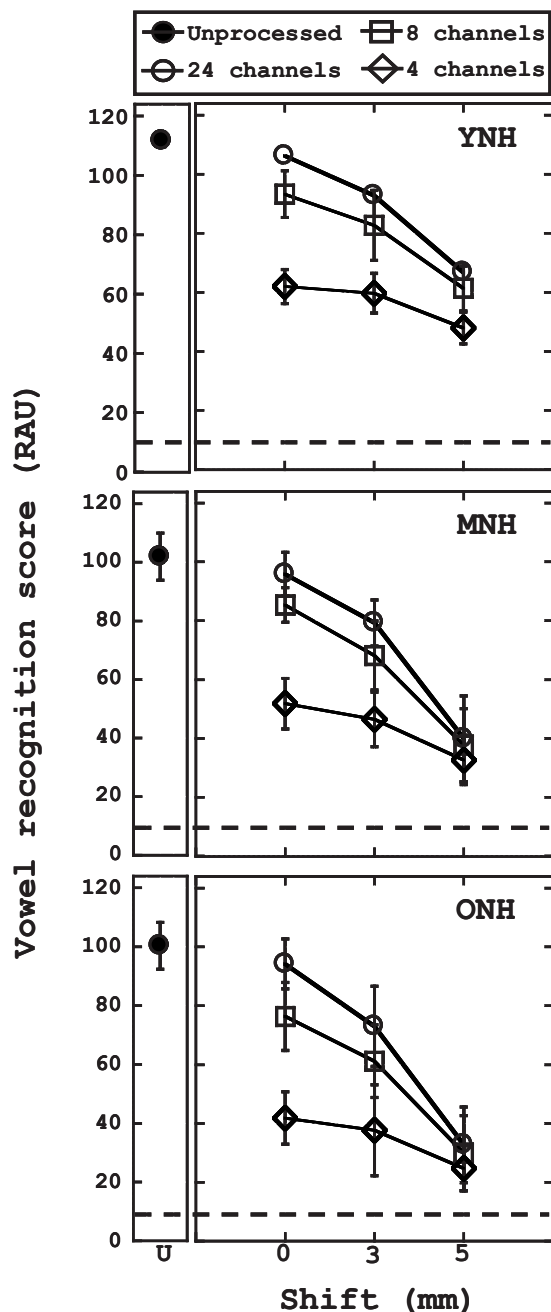


FIG. 3. Vowel recognition (shifted) as a function of the number of channels and degree of spectral shift (in mm). Results for the unprocessed stimuli (U) are also shown for reference. Recognition scores (RAUs) are plotted along the ordinate. Each panel represents results obtained in each age group. The dashed line represents chance performance (8.3%, 6.2 RAUs). Error bars represent  $\pm 1$  standard deviation from the mean score within each condition.

crease in performance as the stimuli were shifted from 0 to 5 mm (collapsed across the number of channels) is dependent on the age of the listener. That is, the performance of the YNH listeners decreased by an average of 28.51 RAUs as stimuli were shifted from 0 to 5 mm, whereas the performance of MNH and ONH listeners decreased by an average of 40.91 and 41.67 RAUs, respectively. Statistical analysis showed that this average decline in performance was less for the YNH listeners, compared to the other two age groups [ $F(2,29)=9.872, p<0.05$ ].

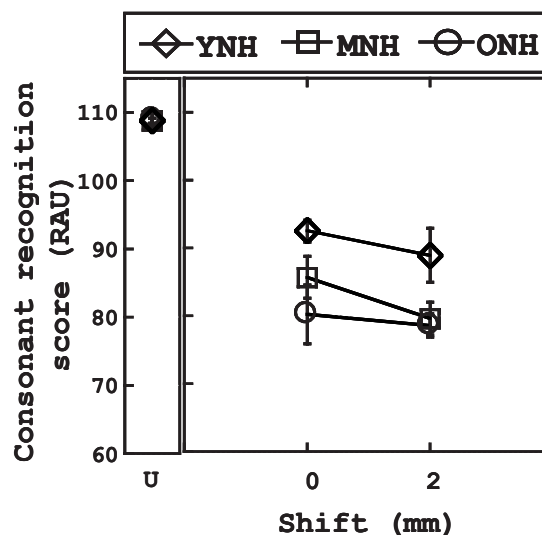


FIG. 4. Stop-consonant recognition (shifted) as a function of the number of channels and degree of spectral shift (in mm). Results for the unprocessed stimuli (U) are also shown as reference. Recognition scores (RAUs) are plotted along the ordinate. Chance performance was 16.7% (19.02 RAUs). Error bars represent  $\pm 1$  standard deviation from the mean score within each group.

### 3. Stop-consonant identification, with spectral shifting

Small but significant effects of age were observed in the results obtained with stop consonants. The results of the stop-consonant identification are shown in Fig. 4. The abscissa indicates the degree of spectral shift (in mm), while the ordinate is the average RAU score. Results with unprocessed stimuli are also shown for reference. A mixed ANOVA was conducted with one within-subject factor [shift (two levels: 0, 2 mm)] and one between-subject factor [group (three levels)]. Results revealed a significant main effect of group [ $F(2,27)=4.36, p<0.05$ ]. *Post hoc* testing (Tukey HSD) revealed that the performance of the YNH listeners was significantly better than that of the ONH listeners ( $p<0.05$ ). Results also reveal that a significant main effect of shift is present [ $F(1,27)=7.52, p<0.05$ ], as the 2 mm shifted condition resulted in slightly worse performance (82.83 RAUs) than the unshifted condition (85.54 RAUs). The group  $\times$  shift interaction was not significant ( $p<0.05$ ).

### B. Factors related to shifted phoneme recognition performance

Much of the data thus far suggest that the YNH group outperformed both the MNH and ONH groups when listening to spectrally degraded and shifted vowels. However, significant variability in performance was noted across all three groups, particularly among those in the MNH and ONH groups. Furthermore, there were slight age differences on the shifted stop-consonant performance. We used a hierarchical multiple-regression analysis to determine how several factors, including both peripheral and cognitive factors, might explain the additional variance of performance on the shifted phoneme recognition tasks. For example, it was thought that perhaps disparate high-frequency thresholds present between the YNH and older (MNH, ONH) groups or differences in

TABLE IV. Descriptive statistics for all of the possible factors considered for the multiple-regression analyses. Minimum and maximum scores are provided, along with average scores and corresponding standard deviations (SDs). Data are shown within each age group (YNH, MNH, and ONH). Units for each of the variables are listed below the name of each factor in parentheses. For the cognitive measures, both raw scores and z-scores are provided. Averaged shifted phoneme scores are shown in RAUs only.

	YNH		MNH		ONH	
	Min-max	Average (SD)	Min-max	Average (SD)	Min-max	Average (SD)
HFPTA (dB HL)	2.5–10.83	6.24 (2.99)	6.66–22.5	14.74 (4.81)	4.16–20	14.66 (5.31)
AGE (years)	21–25	22.8 (1.68)	41–50	45.9 (2.88)	61–71	65.2 (4.04)
FDS (raw score)	11–16	12.8 (1.68)	9–16	11.8 (1.93)	8–15	11 (2.35)
FDS (z-score)	–0.42–1.98	0.448 (0.81)	–1.38–1.98	–0.02 (0.93)	–1.85–1.50	–0.42 (0.55)
BDS (raw score)	6–13	9.80 (2.34)	6–14	8.5 (2.67)	4–12	7.7 (2.71)
BDS (z-score)	–1.01–1.64	0.428 (0.88)	–1.0–1.63	0.45 (0.89)	–1.76–1.26	–0.36 (1.02)
LNS (raw score)	11–20	14.9 (2.92)	9–17	12.5 (2.63)	8–14	10.7 (2.05)
LNS (z-score)	–0.56–2.40	0.726 (0.96)	–1.22–1.41	–0.06 (0.87)	–1.55–0.42	–0.66 (0.68)
SEARCH (raw score)	35–59	45 (7.94)	23–46	36.6 (7.27)	19–48	31.5 (8.65)
SEARCH (z-score)	–0.27–2.35	0.791 (0.80)	–1.57–0.93	–0.09 (0.79)	–2.0–0.33	–0.7 (0.83)
CODE (raw score)	59–104	91.7 (13.47)	63–94	77.9 (11.48)	42–72	62.9 (8.94)
CODE (z-score)	–1.13–1.62	0.871 (0.83)	–0.09–1.01	0.02 (0.7)	–2.17–0.34	–0.90 (0.55)
AVG MEM (z-score)	–0.66–1.78	0.534 (0.78)	–1.07–1.8	–0.05 (0.76)	–1.34–0.69	–0.48 (–0.8)
AVG SPEED (z-score)	–0.7–1.87	0.82 (0.75)	–1.23–0.97	–0.03 (0.66)	–2.09–0.17	–0.8 (–0.62)
3–5 AVG (RAUs)	52.28–78.79	68.66 (7.75)	36.02–65.98	50.7 (9.6)	30.48–61.52	43.23 (11.22)
2 MM (RAUs)	67.95–104.95	88.94 (11.8)	74.26–87.12	79.48 (4.02)	57.24–102.6	78.57 (11.83)

top-down cognitive processes could have contributed to performance. The possible factors considered were average audiometric high-frequency thresholds from 4000–8000 Hz (HF PTA), age of the listener (AGE), SEARCH score, CODE score, FDS score, BDS score, and LNS score. Raw scores obtained on all cognitive measures were converted to z-scores prior to analysis. The dependent variable for the first two regression analyses was average performance on the 3 and 5 mm shifted vowel recognition (3–5 AVG) task (in RAUs), as individual listener performance in these conditions was highly intercorrelated ( $r=0.919$ ,  $p<0.001$ ). The dependent variable on the last regression analysis was average performance on the 2 mm shifted stop-consonant condition (2 MM). Descriptive statistics (minimum, maximum, and average scores) for all factors are provided in Table IV. Scores are reported within each group, and the corresponding units for each factor are provided below in parentheses. Performances for the cognitive measures are provided in both raw scores and z-scores.

### 1. Regression analysis for shifted vowel recognition across all groups

Prior to the regression analysis, the possible predictor variables and dependent variable were entered into a correlational analysis to examine multicollinearity between variables. Although several predictors were intercorrelated, none was *highly* intercorrelated ( $r>0.90$ ). The resulting correlation matrix is shown in Table V; all variables were significantly correlated with the dependent variable 3–5 AVG. Because the current study included a limited number of participants, we thought it prudent to reduce the number of predictors entered in the analysis. Therefore, the z-scores for the three measures of verbal memory span and working memory (FDS, BDS, and LNS) were averaged to create a single predictor variable, MEM AVG. Similarly, the z-scores for the two measures of speed of processing (CODE, SEARCH) were averaged to create a single predictor variable, SPEED AVG.

The results of the ANOVA suggest that performance on the shifted vowel recognition task is strongly related to the

TABLE V. Correlations between all possible predictor variables and two dependent variables: 3–5 AVG and 2 MM.

	2 MM	AGE	HF PTA	FDS	BDS	LNS	SEARCH	CODE
3–5 AVG	0.603 <sup>a</sup>	-0.776 <sup>a</sup>	-0.397 <sup>a</sup>	0.536 <sup>a</sup>	0.531 <sup>a</sup>	0.610 <sup>a</sup>	0.666 <sup>a</sup>	0.708 <sup>a</sup>
2 MM		-0.459 <sup>b</sup>	-0.174	0.550 <sup>a</sup>	0.312	0.496 <sup>a</sup>	0.428 <sup>b</sup>	0.263
AGE			0.616 <sup>a</sup>	-0.407 <sup>b</sup>	-0.349	0.576 <sup>a</sup>	0.684 <sup>a</sup>	-0.782 <sup>a</sup>
HF PTA				-0.214	-0.151	-0.389 <sup>b</sup>	-0.401 <sup>b</sup>	-0.426 <sup>b</sup>
FDS					0.644 <sup>a</sup>	0.721 <sup>a</sup>	0.413 <sup>b</sup>	0.169
BDS						0.482 <sup>a</sup>	0.332	0.141
LNS							0.546 <sup>a</sup>	0.494 <sup>a</sup>
SEARCH								0.778 <sup>a</sup>

<sup>a</sup> $p < 0.01$ .

<sup>b</sup> $p < 0.05$ .

age of the listener (see Fig. 3). Not surprisingly, a strong correlation is also present between these two variables ( $r = -0.776$ ,  $p < 0.01$ ), as shown in Table V. Furthermore, stepwise regression analysis was performed for exploratory purposes and the correlation matrix was examined to determine which variables should be entered into the regression analysis. Based on these observations, it was determined that AGE, MEM AVG, SPEED AVG, and HF PTA should be selected as the four predictors entered into the regression analysis. All analyses suggested that AGE was the most significant factor related to 3–5 AVG, and therefore it was entered into the first block of the hierarchical regression analysis. Next, the factor of HF PTA was entered in the second block, and then the two cognitive measures (MEM AVG and SPEED AVG) were entered in the third and final block. The criterion used for entering a variable into the final model was  $p < 0.05$ .

This regression is shown in Fig. 5, with the predicted values (abscissa) plotted against the observed vowel recognition data (ordinate). Data points are coded by age group.

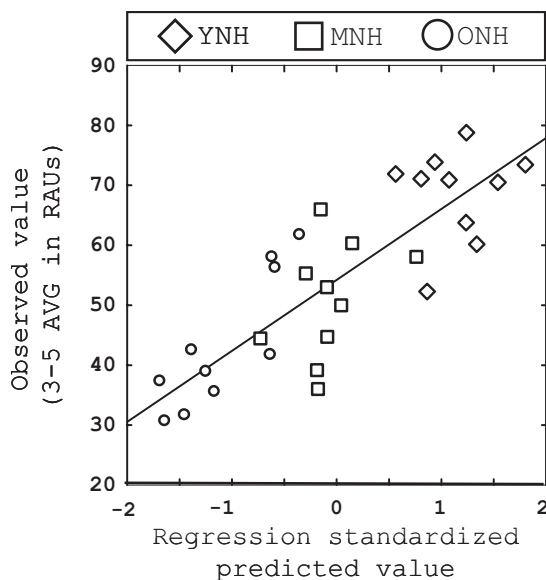


FIG. 5. Linear regression for shifted vowel recognition when all three age groups (YNH, MNH, ONH) were included in the analysis. The observed value (average scores for 3 and 5 mm shifts in RAUs) is plotted against the standardized predicted value for each participant. Data points are coded according to each age group.

As expected, the factor of AGE was a strong predictor of performance ( $R^2 = 0.601$ ,  $p < 0.001$ ), while MEM AVG added slightly but significantly resulting in a final  $R^2 = 0.686$  ( $p < 0.001$ ), and the final regression equation:

$$3 - 5 \text{ AVG} = 75.664 - 0.481(\text{AGE}) + 5.617(\text{MEM AVG}). \quad (1)$$

These results suggest that while the age of the listener is an important factor in predicting performance when listening to noise-band vocoded vowels, verbal memory capacity also contributes, albeit to a smaller extent. Speed of processing (SPEED AVG) and high-frequency hearing sensitivity (HF PTA) were not significant predictors of performance.

## 2. Regression analysis for shifted vowel recognition within middle-aged and older groups

Visual inspection of the results obtained with shifted vowels (Fig. 3) shows that while the YNH group undoubtedly performed better than the MNH and ONH groups, performance across the two latter groups was more comparable. Furthermore, significant variability was noted within each age group as some ONH listeners performed exceptionally well, and some MNH listeners performed exceptionally poorly, suggesting that age may not be the main factor influencing performance within these two groups. It was of interest, therefore, to examine the influence of other factors in more details among these two older groups as a possible source to explain the variance in performance.

The same predictor variables mentioned previously were tested for multicollinearity and significance. Results from the two groups (MNH and ONH) are shown in Table VI and indicated that all of the predictor variables were significant predictors of 3–5 AVG with the exception of HF PTA. Cognitive measures were somewhat interrelated but not highly correlated with one another. Therefore, similar to the previous regression analysis, the combined variables of MEM AVG and SPEED AVG were considered for entry into the analysis. Preliminary stepwise multiple-regression analyses were performed, along with examination of the correlational matrix, to determine which variables should be entered into the final analysis and to determine the order in which each variable should be entered into the hierarchical regression analysis. Based on these results, it was determined that the

TABLE VI. Identical to Table V, but within MNH and ONH groups alone.

	AGE	HF PTA	FDS	BDS	LNS	SEARCH	CODE
3-5 AVG	-0.472 <sup>a</sup>	0.170	0.536 <sup>b</sup>	0.519 <sup>a</sup>	0.493 <sup>a</sup>	0.628 <sup>b</sup>	0.558 <sup>a</sup>
AGE		0.019	-0.292	-0.224	-0.394	-0.582 <sup>b</sup>	-0.718 <sup>b</sup>
HF PTA			0.111	0.083	-0.007	-0.039	0.007
FDS				0.651 <sup>b</sup>	0.667 <sup>b</sup>	0.411	0.027
BDS					0.246	0.188	-0.127
LNS						0.400	0.363
SEARCH							0.638 <sup>b</sup>

<sup>a</sup> $p < 0.05$ .

<sup>b</sup> $p < 0.01$ .

cognitive factors (MEM AVG and SPEED AVG) were better predictors of performance than the factor of AGE. Therefore, in the hierarchical analysis, we entered MEM AVG and SPEED AVG into the first block, followed by AGE in the second block. The criterion used for entering a variable into the final model was  $F < 0.05$ .

The regression line is shown in Fig. 6 with the predicted values (abscissa) plotted against the observed vowel recognition data (ordinate). Data points are coded by age group. The resulting multiple-regression analyses revealed that MEM AVG accounted for a moderate portion of the variance ( $R^2 = 0.434$ ,  $p < 0.01$ ), while a second variable (SPEED AVG) also added slightly but significantly (change in  $R^2 = 0.246$ ,  $p < 0.01$ ) to the final regression model. The factor of AGE did not significantly contribute to the model ( $t = 0.873$ ,  $p > 0.05$ ) above and beyond the variance already accounted for by the cognitive factors, and therefore was not included in the final model. The final regression model accounted for a significant proportion of the variance noted among the MNH and ONH participants ( $R^2 = 0.680$ ,  $p < 0.01$ ). The final regression equation was

$$3 - 5 \text{ AVG} = 52.036 + 7.135(\text{MEM AVG}) + 7.601(\text{SPEED AVG}). \quad (2)$$

These results suggest that when a regression analysis is performed within the MNH and ONH groups only, cognitive factors (MEM AVG and SPEED AVG) account for the most variance among performance on 3-5 AVG. In contrast with the first regression analysis (all groups), AGE was not the foremost factor related to shifted vowel recognition. Furthermore, SPEED AVG was significantly related to performance within the MNH and ONH groups; this is also in contrast with the first regression analysis including all three age groups.

### 3. Regression analysis for shifted stop-consonant recognition across all three groups

Similar to the previous analyses, a third regression analysis was performed to assess the factors that influence shifted stop-consonant performance (2 MM). Examination of the correlation matrix (see Table V) and exploratory stepwise regression analyses showed that MEM AVG was the strongest predictor of performance, and therefore was entered into the first block of the hierarchical regression analysis. The factor of AGE was entered into the second block, and SPEED AVG was entered into the third block. It should be noted that HF PTA was not a significant predictor of performance (2 MM) and therefore was not entered into the regression analysis. The criterion used for entering a variable into the final model was  $F < 0.05$ .

Results are displayed in Fig. 7 and show that MEM AVG accounted for a small but significant portion of the variance ( $R^2 = 0.276$ ,  $p < 0.05$ ) on shifted stop-consonant recognition performance (2 MM). Both of the remaining factors (SPEED AVG and AGE) failed to make significant contributions to the model. The final model was

$$2 \text{ MM} = 82.331 + 6.518(\text{MEM AVG}). \quad (3)$$

Although the results of all regression analyses appear to be meaningful, they should be interpreted with considerable caution due to the limited number of participants (20 or 30) and total number of variables (3 or 4) entered into the equation. Nonetheless, these results would suggest that cognitive abilities play a role in the immediate perception of spectrally

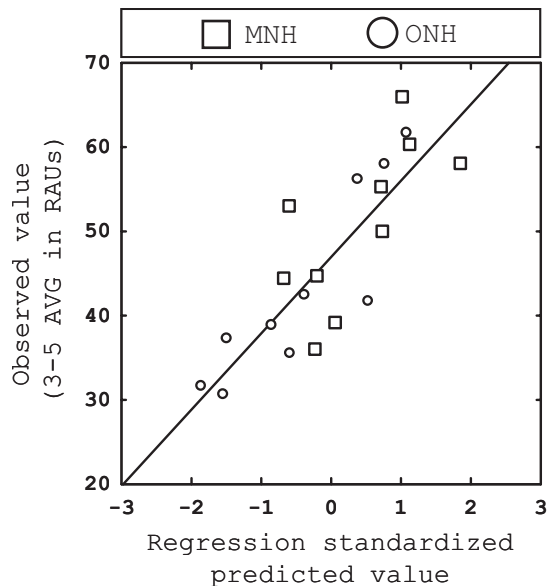


FIG. 6. Linear regression for shifted vowel recognition when only the MNH and ONH groups were included in the analysis. The observed value (average scores for 3 and 5 mm shifts in RAUs) is plotted against the standardized predicted value for each participant. Data points are coded according to each age group.



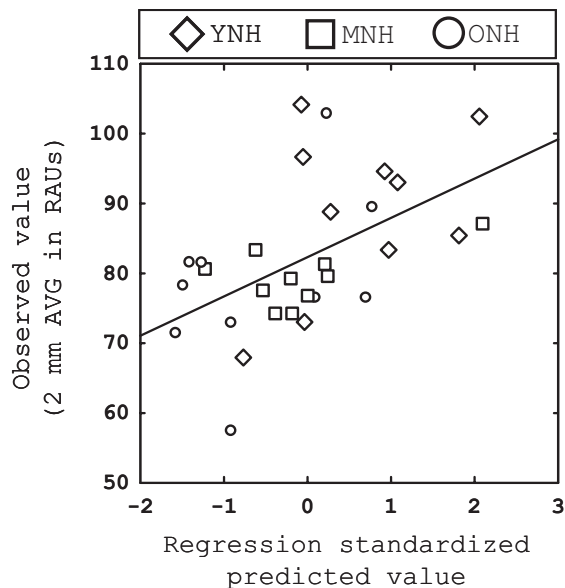


FIG. 7. Linear regression for shifted stop consonants when all three age groups (YNH, MNH, ONH) were included in the analysis. The observed value (average scores for the 2 mm shifted condition in RAUs) is plotted against the standardized predicted value for each participant. Data points are coded according to each age group.

degraded vowels among adult listeners and help to explain the additional variance noted among scores, which cannot be accounted for by age alone.

#### IV. DISCUSSION

##### A. Summary of results and comparison with previous investigations

It is obvious that spectral degradation of phonemes results in poor recognition among adult listeners of various ages. In many aspects, results from the current investigation are similar to those reported previously. For example, the recognition of consonants is generally more robust to spectral distortion compared to that of vowels (Dorman *et al.*, 1997; Baskent and Shannon, 2003), and this pattern persisted in the current investigation regardless of the age of the listener. Previous findings among adult listeners have shown a consistent relationship between spectral degradation and phoneme recognition performance, as performance tends to decline steadily as the number of channels decreases or as the frequency-to-place mismatch increases (Fishman *et al.*, 1997; Fu and Shannon, 1999; Friesen *et al.*, 2001; Baskent and Shannon, 2003). However, many of the previous studies measuring this relationship did not control for any contributions of age. Based on the findings of Souza and colleagues (Souza and Kitch, 2001; Souza and Boike, 2006) as well as those reported by Sheldon *et al.* (2008), one would expect that some differences would be present between the performance of younger and older adult listeners when they are required to listen through a limited number of spectral channels and when feedback is not provided. Similar results were also evident in the current investigation. In addition, significant differences in performance were present across the three age groups when a spectral shift was added to mimic the limited CI insertion depth. A lack of a significant group

× channel interaction for both the unshifted vowel and consonant stimuli suggests that overall performance was worse for the ONH group compared to the YNH group, and that this relationship was not dependent on the number of channels. This finding is apparent from examining the data shown in Fig. 1 and 2. Although significant age differences are present as the spectral resolution is decreased, similar age differences are also present in the unprocessed condition. Therefore, all age groups appear to be equally robust to spectral distortion when performing vowel and consonant recognition tasks, at least when stimuli are *not* spectrally shifted. Adding a spectral shift creates further distortion and exposes differential performance across the three age groups.

These findings differ from those reported by Peelle and Wingfield (2005), who reported little or no age-related differences when they presented spectrally degraded and compressed sentences to younger and older adults. Although the reason for this is not completely clear, it is likely that the choice of stimuli (sentences) provided more contextual cues compared to the phoneme stimuli used in the current study. Contextual cues are used by older listeners to derive meaning and improve perception of the stimulus, as several studies report improvements in speech understanding among older adults when highly contextual stimuli are presented, compared to lower contextual stimuli (Pichora-Fuller *et al.*, 1995; Gordon-Salant and Fitzgibbons, 1997). Furthermore, the methodology used by Peelle and Wingfield (2005), who applied an apical shift to their stimuli, was not intended to closely mimic that of CI processing which results in a basal, not apical, shift. Studies have shown that sentence and phoneme recognition are negatively, but unequally, affected by compressive, expansive, and ordinary shifted spectral distortion (Baskent and Shannon, 2003; 2004; 2005). For these reasons, differences in methodology and stimuli make any direct comparisons between the current study and Peelle and Wingfield (2005) rather difficult.

The severe reduction in spectral resolution inherent to noise-band vocoded stimuli requires participants to rely more on temporal cues within the signal comparable to what is required of CI users. Temporal deficits have been reported among middle-aged individuals, and are pervasive among older NH and HI listeners; it is possible that underlying temporal processing deficits contributed to the decreased performance among the MNH and ONH listeners in the current study. However, further investigation is needed as this hypothesis cannot be confirmed based on our results.

##### B. Estimating performance with cochlear-implants

The results of the current study suggest that a spectral shift of 3 or 5 mm can affect vowel recognition performance of middle-aged and older listeners, more so than that of younger listeners. Imaging studies demonstrate that the insertion depth of the electrode array can range anywhere from 14–25 mm (approximately), with an average insertion depth of 20 mm (Ketten *et al.*, 1998; Skinner *et al.*, 2002). Applying Greenwood's equation (Greenwood, 1990), the most apical electrode would be at a place in the cochlea primed to receive anything from a 3000 Hz signal (14 mm insertion

depth) to a 500 Hz signal (25 mm insertion depth), with an average insertion depth of 1000 Hz (20 mm). Modern CI processors filter the incoming speech signal using a set of bandpass filters, the lowest of which typically has a frequency of 200–300 Hz. This means that even with optimal placement (25 mm insertion depth), the incoming 300 Hz signal is being processed at a place in the cochlea that is primed to receive 500 Hz information resulting in at least a 3 mm shift or nearly an octave. It is likely that CI users are listening through eight to ten spectral channels, at best (Fishman *et al.*, 1997; Friesen *et al.*, 2001). By applying these parameters to the results of the current investigation, it is hypothesized that even under optimal circumstances the age of the adult listener will have an effect on immediate vowel recognition performance. The data presented in Fig. 3 show that with eight channels of information and a 3 mm shift, disparities are present across the three age groups: the average performances of the three groups are 82.67, 68.22, and 61.20 RAUs for YNH, MNH and ONH, respectively. The disparities between the groups' performance become larger as spectral distortion is increased (number of channels decreased, spectral shift increased) and, in particular, the performance of the YNH group was more robust to spectral distortion compared to the MNH and ONH listeners.

### C. Regression analyses and associations with cognitive measures

Results from the first regression analysis (all groups) suggest that age accounts for most of the variance noted among the shifted vowel recognition performance ( $R^2 = 0.601$ ), while average memory score (MEM AVG) only adds slightly (change in  $R^2 = 0.085$ ) to the final regression model ( $R^2 = 0.686$ ). Conversely, when shifted vowel data are analyzed within the MNH and ONH groups only, the factors of memory (MEM AVG) and speed of processing (SPEED AVG) seem to be better predictors of performance than AGE alone. Analysis of the stop-consonant data showed that performance across all three groups was best predicted by memory function (MEM AVG). From a cognitive perspective, it could be inferred that individuals listening through a CI are subject to increased cognitive load, as they are required to derive meaningful speech from a severely degraded stimulus. Cognitive theory supports the notion that such a task would become more difficult as the age of the listener increases (Sweller, 1988; 1994; Van Gerven *et al.*, 2002). In fact, recent studies have shown that cognitive abilities of adult CI listeners influence postimplantation speech recognition performance (Lyxell *et al.*, 1996; Heydebrand *et al.*, 2007). Heydebrand *et al.* (2007) reported that preimplant measures of verbal working memory and learning were significant predictors of speech recognition abilities obtained at six month postactivation in adult CI listeners aged 24–60. Similarly, a wealth of research focusing on pediatric CI listeners provides support that cognitive abilities play an important role in postimplantation speech and language performance (Eisenberg *et al.*, 2000; Pisoni and Geers, 2000; Cleary *et al.*, 2001; Geers, 2003; Geers *et al.*, 2003; Pisoni and Cleary, 2003). The current study does contrast with pre-

vious studies on one main point: past investigations have tied verbal memory (working memory in particular) capacity to actual *speech* or discourse comprehension (ie: sentences); it is commonly believed that phoneme recognition is not as reliant on memory capacities (see Pichora-Fuller, 2003 for review). It is proposed, however, that when phonemes are severely degraded (as in the current study), a different pattern likely emerges. A severely degraded phoneme lacks both acoustic and contextual cues and therefore requires significant top-down processing in order that it be correctly identified. Therefore, this type of task might be more related to cognitive abilities compared to one in which phonemic stimuli are undistorted. Furthermore, a different pattern could be evident when severely degraded, highly contextual sentences are used as stimuli; however this hypothesis would require additional investigation.

Results from the regression analyses should be interpreted with caution due to the limited number of participants in the current study and the limited upper age range of the older participants. It could be argued that the partial cognitive test battery used in the current study is also a limiting factor, as memory and speed of processing abilities may not be the best, or only cognitive measures related to performance. Our hypothesis was that chronological age was not the *only* contributing factor when understanding severely spectrally degraded phonemes, and it seemed logical that higher order top-down processes would also play a role. However, it remains unknown how to best capture such top-down processing abilities. We chose to measure abilities in two well-known cognitive domains (memory and speed of processing) in an attempt to assess general cognitive abilities among our listeners; the fact that we did find significant relationships between such cognitive abilities and perceptual measures does not suggest direct causation. Furthermore, we are not proposing that performance on these diverse tasks (shifted vowel recognition, memory abilities, and speed of processing abilities) are necessarily determined by the same underlying cause [“common cause hypothesis” Lindenberger and Baltes, 1994]. Rather, the strong correlation observed in our results simply lends support to the notion that cognitive function is predictive of speech recognition by middle-aged and older listeners in adverse conditions.

It is believed that measures of working memory capture some fundamental ability central to several measures of cognition, including fluid intelligence and language abilities (Conway *et al.*, 2002; Kemper *et al.*, 2003) making it a valid and compelling measure. Although the current study did not investigate performance among actual CI users, the results from the current study are in agreement with previous studies that report significant relationships between working memory ability and outcomes with cochlear implantation. The contribution of speed of processing on the ability to comprehend spectrally degraded vowels is less clear, as this measure is theoretically associated with tasks measuring speed of execution or ability to comprehend rapid speech (Salthouse, 1996, 2000). Nonetheless, speed of processing ability did account for some of the variances noted among performance and is therefore noteworthy; further investigation is needed to validate such findings.

## D. Potential interactions between aging, training and aural rehabilitation

Outcomes of the current investigation reflect the immediate untrained performance of a limited number of individuals. Emerging literature strongly conveys the importance of training and adaptation when listening to spectrally degraded stimuli, as most CI listeners (and those NH individuals listening to noise-band vocoded stimuli) are able to adapt extremely well over time (Rosen *et al.*, 1999; Svirksy *et al.*, 2004; Fu *et al.*, 2002; 2005; see Fu and Galvin, 2007 for review). Furthermore, the age effects found in the current study using phonemic stimuli do not necessarily reflect or predict performance in real-world environments, in which highly contextual conversations dominate the listener's everyday experience. Nonetheless, the results from the current study do suggest that age might influence perception of spectrally degraded stimuli.

Although rehabilitation is commonly mandated for pediatric cochlear implantees, rehabilitation of the adult CI recipient is not a standard practice as most users are able to adapt adequately without structured rehabilitation services. Results from Fu and Galvin (2007) would suggest that even adequate CI users can improve their performance when provided with aural rehabilitation. Further evidence suggests that older listeners receive comparable benefit when trained on difficult speech recognition tasks, when compared to their younger counterparts (Peelle and Wingfield, 2005; Burk *et al.*, 2006; Golomb *et al.*, 2007); this implies that CI recipients of all ages might be capable of improving their performance with structured rehabilitation. The results of the current investigation imply that older listeners are at a significant deficit for accurately recognizing CI processed speech. An aural rehabilitation program may be particularly beneficial to older adult CI users.

## V. CONCLUSIONS

(1) The age of the adult listener influences the immediate perception of spectrally degraded phonemes. In particular, performance of the YNH group exceeded that of the MNH and ONH listeners. In this study, the largest effect was seen on vowel recognition tasks that were processed to include a limited number of spectral bands and to simulate a frequency-to-place mismatch. Conversely, performance among the MNH and ONH listeners was quite comparable, with ONH listeners performing only slightly worse than those listeners in the MNH group.

(2) Results from the hierarchical multiple-regression analysis suggest that chronological age is the primary predictor of shifted vowel recognition performance when data are analyzed across all three age groups (YNH, MNH, and ONH). Verbal memory abilities also contribute slightly, but significantly, to performance. Conversely, the regression analysis within the MNH and ONH groups alone suggest that cognitive abilities are a better predictor of performance than chronological age.

(3) It appears that the immediate perception of spectrally degraded phonemes among adult listeners is related to both chronological age and cognitive factors; however, the degree

to which each variable exerts influence on performance is dependent on the age of the listener. Performance among the YNH listeners was fairly independent of cognitive abilities, whereas performance among the MNH and ONH listeners was significantly related to cognitive abilities.

(4) The results from this study have important clinical implications regarding the rehabilitation and counseling of adult CI recipients. Further research is warranted to examine how chronological age and cognitive abilities influence adaptation to spectrally degraded and shifted speech over an extended period of time. Furthermore, future research should focus on an in-depth analysis of the relationship between various cognitive measures and understanding severely spectrally degraded stimuli to better understand the interaction between bottom-up and top-down processes on this type of task.

## ACKNOWLEDGEMENTS

We would like to thank the participants for their time and willingness to contribute to this study. We are grateful to Dr. Qian-Jie Fu for the use of the software used for the phoneme recognition task and to John J. Galvin III for his helpful insight when designing this project. We thank two anonymous reviewers and the Associate Editor for their comments on an earlier version of the manuscript. This work was funded by NIDCD Grant No. R01-DC004786 to M.C. K.C.S. was supported in part by training Grant No. T32-DC00046 from NIDCD.

<sup>1</sup>Two ONH participants did not fit these criteria used to assess normal middle-ear function. ONH-3 had negative middle ear pressure that fell just outside normal limits (-105 daPa, left ear). Acoustic reflex thresholds could not be measured in ONH-10 because a hermetic seal could not be maintained. Click-evoked transient otoacoustic emissions were present and robust across both ears in these two subjects, suggesting a lack of significant middle ear pathology.

- ANSI (2004). "American National Standard Specification for Audiometers," ANSI S3.6-2004 (American National Standards Institute, New York).
- Baskent, D., and Shannon, R. V. (2005). "Interactions between cochlear implant electrode insertion depth and frequency-place mapping," *J. Acoust. Soc. Am.* **117**, 1405-1416.
- Baskent, D., and Shannon, R. V. (2004). "Frequency-place compression and expansion in cochlear implant listeners," *J. Acoust. Soc. Am.* **116**, 3130-3140.
- Baskent, D., and Shannon, R. V. (2003). "Speech recognition under conditions of frequency-place compression and expansion," *J. Acoust. Soc. Am.* **113**, 2064-2076.
- Blamey, P., Arndt, P., Bergeron, F., Bredberg, G., Brimacombe, J., Facer, G., Larky, J., Lindstrom, B., Nedzelski, J., Peterson, A., Shipp, D., Staller, S., and Whitford, L. (1996). "Factors affecting auditory performance of post-linguistically deaf adults using cochlear implants," *Audiol. Neuro-Otol.* **1**, 293-306.
- Bodmer, D., Shipp, D. B., Ostroff, J. M., Ng, A. H. C., Stewart, S., Chen, J. M., and Nedzelski, J. M. (2007). "A comparison of postcochlear implantation speech scores in an adult population," *Laryngoscope* **117**, 1408-1411.
- Burk, M. H., Humes, L. E., Amos, N. E., and Stauser, L. E. (2006). "Effect of training on word-recognition performance in noise for young normal-hearing and older-hearing impaired listeners," *Ear Hear.* **27**, 263-278.
- Clark, G. (2003). *Cochlear Implants: Fundamentals and Applications* (Springer-Verlag, New York).
- Cleary, M., Pisoni, D. B., and Geers, A. E. (2001). "Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants," *Ear Hear.* **22**, 395-411.
- Conway, A. R. A., Cowan, N., Buting, M. F., Theriault, D. J., and Minkoff,



- S. R. B. (2002). "A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence," *Intelligence* **30**, 163–183.
- Dorman, M. F., Loizou, P. C., Fitzke, J., and Tu, Z. (1998). "The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6–20 channels," *J. Acoust. Soc. Am.* **104**, 3583–3585.
- Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs," *J. Acoust. Soc. Am.* **102**, 2403–2411.
- Dorman, M., Dankowski, K., McCandless, G., Parkin, J. L., and Smith, L. (1990). "Longitudinal changes in word recognition by patients who use the ineraid cochlear implant," *Ear Hear.* **11**, 455–459.
- Dorman, M., Hannley, M., Dankowski, K., Smith, L., and McCandless, G. (1989). "Word recognition by 50 patients fitted with the Symbion multi-channel cochlear implant," *Ear Hear.* **10**, 44–49.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2002). "Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing," *J. Acoust. Soc. Am.* **111**, 2897–2907.
- Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., and Boothroyd, A. (2000). "Speech recognition with reduced spectral cues as a function of age," *J. Acoust. Soc. Am.* **107**, 2704–2710.
- Facer, G. W., Peterson, A. M., and Brey, R. H. (1995). "Cochlear implantation in the senior citizen age group using the Nucleus 22-channel device," *Ann. Otol. Rhinol. Laryngol. Suppl.* **166**, 187–190.
- Fishman, K. E., Shannon, R. V., and Slattery, W. H. (1997). "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," *J. Speech Lang. Hear. Res.* **40**, 1201–1215.
- Fitzgibbons, P. G., Gordon-Salant, S., and Barrett, J. (2007). "Age-related differences in discrimination of an interval separating onsets of successive tone bursts as a function of interval duration," *J. Acoust. Soc. Am.* **122**, 458–466.
- Fitzgibbons, P. J., and Gordon-Salant, S. (2004). "Age effects on discrimination of timing in auditory sequences," *J. Acoust. Soc. Am.* **116**, 1126–1134.
- Fitzgibbons, P. J., and Gordon-Salant, S. (2001). "Aging and temporal discrimination in auditory sequences," *J. Acoust. Soc. Am.* **109**, 2955–2963.
- Folstein, M., Folstein, S., and McHugh, P. (1975). "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician," *J. Psychiatr. Res.* **12**, 189–198.
- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). "Speech recognition in noise as the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* **110**, 1150–1163.
- Fu, Q.-J., and Galvin, J. J. (2007). "Perceptual learning and auditory training in cochlear implant recipients," *Trends Amplif.* **11**, 195–205.
- Fu, Q.-J., Nogaki, G., and Galvin, J. J. (2005). "Auditory training with spectrally shifted speech: Implications for cochlear implant patient auditory rehabilitation," *J. Assoc. Res. Otolaryngol.* **180**, 180–190.
- Fu, Q.-J., Shannon, R. V., and Galvin, J. J. (2002). "Perceptual learning following changes in the frequency-to electrode assignment with the Nucleus-22 cochlear implant," *J. Acoust. Soc. Am.* **112**, 1664–1674.
- Fu, Q.-J., and Shannon, R. V. (1999). "Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing," *J. Acoust. Soc. Am.* **105**, 1889–1900.
- Fu, Q.-J., Shannon, R. V., and Wang, X. (1998). "Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing," *J. Acoust. Soc. Am.* **104**, 3586–3596.
- Gantz, B. J., Woodworth, G. G., Knutson, J. F., Abbas, P. J., and Tyler, R. S. (1993). "Multivariate predictors of audiological success with multichannel cochlear implants," *Ann. Otol. Rhinol. Laryngol.* **102**, 909–916.
- Geers, A. E. (2003). "Predictors of reading skill development in children with early cochlear implantation," *Ear Hear.* **24**, 59S–68S.
- Geers, A. E., Nicholas, J. G., and Sedey, A. L. (2003). "Language skills of children with early cochlear implantation," *Ear Hear.* **24**, 46S–58S.
- Gelfand, S. A., Schwander, T., and Silman, S. (1990). "Acoustic reflex thresholds in normal hearing and cochlear impaired ears: Effects of no-response rates on 90th percentiles in a large sample," *J. Speech Hear. Disord.* **55**, 198–205.
- Golomb, J. D., Peelle, J. E., and Wingfield, A. (2007). "Effects of stimulus variability and adult aging on adaptation to time-compressed speech," *J. Acoust. Soc. Am.* **121**, 1701–1708.
- Gordon-Salant, S., and Fitzgibbons, P. J. (2004). "Effects of stimulus and noise rate variability on speech perception by younger and older adults," *J. Acoust. Soc. Am.* **115**, 1808–1817.
- Gordon-Salant, S., and Fitzgibbons, P. J. (1997). "Selected cognitive factors and speech recognition performance among young and elderly listeners," *J. Speech Lang. Hear. Res.* **40**, 423–431.
- Gordon-Salant, S., and Fitzgibbons, P. J. (1993). "Temporal factors and speech recognition performance in young and elderly listeners," *J. Speech Hear. Res.* **36**, 1276–1285.
- Green, K. M., Bhatt, Y. M., Mawman, D. J., O'Driscoll, M. P., Saeed, S. R., Ramsden, R. T., and Green, M. W. (2007). "Predictors of audiological outcome following cochlear implantation in adults," *Cochlear Implants Int.* **8**, 1–11.
- Greenwood, D. D. (1990). "A cochlear frequency-position function for several species—29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.
- Grose, J. H., Hall, J. W., and Buss, E. (2006). "Temporal processing deficits in the presenescent auditory system," *J. Acoust. Soc. Am.* **119**, 2305–2315.
- Grose, J. H., Hall, J. W., III, and Buss, E. (2001). "Gap duration discrimination in listeners with cochlear hearing loss: Effects of gap and marker duration, frequency separation, and mode of presentation," *J. Assoc. Res. Otolaryngol.* **2**, 388–398.
- Heydebrand, G., Hale, S., Potts, L., Gotter, B., and Skinner, M. (2007). "Cognitive predictors of improvements in adults' spoken word recognition six months after cochlear implant activation," *Audiol. Neuro-Otol.* **12**, 254–264.
- Kelsall, D. C., Shallop, J. K., and Burnelli, T. (1995). "Cochlear implantation in the elderly," *Am. J. Otol.* **16**, 609–615.
- Kemper, S., Herman, R. E., and Lian, C. H. (2003). "The costs of doing two things at once for young and older adults: Talking while walking, finger tapping, and ignoring speech or noise," *Psychol. Aging* **18**, 181–192.
- Ketten, D. R., Skinner, M. W., Wang, G., Vannier, M. W., Gates, G. A., and Neely, J. G. (1998). "In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays," *Ann. Otol. Rhinol. Laryngol. Suppl.* **175**, 1–16.
- Leung, J., Wang, N. Y., Yeagle, J. D., Chinnici, J., Bowditch, S., Francis, H. W., and Niparko, J. K. (2005). "Predictive models for cochlear implantation in elderly candidates," *Arch. Otolaryngol. Head Neck Surg.* **131**, 1049–1054.
- Lindenberger, U., and Baltes, P. B. (1994). "Sensory functioning and intelligence in old age: A strong connection," *Psychol. Aging* **9**, 339–355.
- Lister, J., Besing, J., and Koehnke, J. (2002). "Effects of age and frequency disparity on gap discrimination," *J. Acoust. Soc. Am.* **111**, 2793–2800.
- Lyxell, B., Andersson, J., Arlinger, S., Bredberg, G., Harder, H., and Ronnberg, J. (1996). "Verbal information-processing capabilities and cochlear implants: implications for preoperative predictors of speech understanding," *J. Deaf Stud. Deaf Educ.* **1**, 190–201.
- Nie, K., Barco, A., and Zeng, F.-G. (2006). "Spectral and temporal cues in cochlear implant speech perception," *Ear Hear.* **27**, 208–217.
- Niparko, J. (2004). "Cochlear implants: Clinical applications," in *Cochlear Implants: Auditory Prosthesis and Electric Hearing*, edited by F.-G. Zeng, A. N. Popper, and R. Fay (Spring-Verlag, New York), pp. 53–100.
- Ostroff, J. M., McDonald, K. L., Schneider, B. A., and Alain, C. (2003). "Aging and the processing of sound duration in human auditory cortex," *Hear. Res.* **181**, 1–7.
- Parkin, J. L., Stewart, B. E., Dankowski, K., and Haas, L. J. (1989). "Prognosticating speech performance in multichannel cochlear implant patients," *Otolaryngol.-Head Neck Surg.* **101**, 314–319.
- Pasanisi, E., Bacciu, A., Vincenti, V., Guida, M., Barbot, A., Berghenti, M. T., and Bacciu, S. (2003). "Speech recognition in elderly cochlear implant recipients," *Clin. Otolaryngol.* **28**, 154–157.
- Peelle, J. E., and Wingfield, A. (2005). "Dissociations in perceptual learning revealed by adult age differences in adaptation to time-compressed speech," *J. Exp. Psychol. Hum. Percept. Perform.* **31**, 1315–1330.
- Pichora-Fuller, M. K. (2003). "Cognitive aging and auditory information processing," *Int. J. Audiol.* **42**, 2S26–2S32.
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). "How young and old adults listen to and remember speech in noise," *J. Acoust. Soc. Am.* **97**, 593–608.
- Pisoni, D. B., and Cleary, M. (2003). "Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation," *Ear Hear.* **24**, 106S–120S.
- Pisoni, D. B., and Geers, A. B. (2000). "Working memory in deaf children with cochlear implants: Correlations between digit span and measures of spoken language processing," *Ann. Otol. Rhinol. Laryngol. Suppl.* **185**,



- Rosen, S., Faulkner, A., and Wilson, L. (1999). "Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants," *J. Acoust. Soc. Am.* **106**, 3629–3636.
- Roup, C. M., Wiley, T. L., Safady, S. H., and Stoppenbach, D. T. (1998). "Tymanometric screening norms for adults," *Am J Audiol* **7**, 1–5.
- Rubinstein, J. T., Parkinson, W. S., Tyler, R. S., and Gantz, B. J. (1999). "Residual speech recognition and cochlear implant performance: Effects of implantation criteria," *Am. J. Otol.* **20**, 445–452.
- Salthouse, T. A. (2000). "Aging and measures of processing speed," *Biol. Psychol.* **54**, 35–54.
- Salthouse, T. A. (1996). "The processing-speed theory of adult age differences in cognition," *Psychol. Rev.* **103**, 403–428.
- Schroeder, M. R. (1968). "Reference signal for signal quality studies," *J. Acoust. Soc. Am.* **44**, 1735–1736.
- Shannon, R. V., Feng, F.-G., and Wygonski, J. (1998). "Speech recognition with altered spectral distribution of envelope cues," *J. Acoust. Soc. Am.* **104**, 2467–2476.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Sheldon, S., Pichora-Fuller, M. K., and Schneider, B. A. (2008). "Effects of age, presentation method, and learning on identification of noise-vocoded words," *J. Acoust. Soc. Am.* **123**, 476–488.
- Shin, Y. J., Fraysse, B., Dequine, O., Vales, O., Laborde, M. L., Bouccara, D., Sterkers, O., and Uziel, A. (2000). "Benefits of cochlear implantation in elderly patients," *Otolaryngol.-Head Neck Surg.* **122**, 602–606.
- Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., Neely, J. G., Kletzker, G. R., Brunsten, B., and Blocker, B. (2002). "CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients," *J. Assoc. Res. Otolaryngol.* **3**, 332–350.
- Snell, K. B., and Frisina, D. R. (2000). "Relationships among age-related differences in gap detection and word recognition," *J. Acoust. Soc. Am.* **107**, 1615–1626.
- Souza, P. E., and Boike, K. T. (2006). "Combining temporal-envelope cues across channels: Effects of age and hearing loss," *J. Speech Lang. Hear. Res.* **49**, 138–149.
- Souza, P. E., and Kitch, V. (2001). "The contribution of amplitude envelope cues to sentence identification in young and aged listeners," *Ear Hear.* **22**, 112–119.
- Sterkers, O., Mosnier, I., Ambert-Dahan, E., Herelle-Dupuy, E., Bozorg-Grayeli, A., and Bouccara, D. (2004). "Cochlear implants in elderly people: Preliminary results," *Acta Oto-Laryngol., Suppl.*, 64–67.
- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech Hear. Res.* **28**, 455–462.
- Svirsky, M. A., Silveira, A., Neuburger, H., Teoh, S.-W., and Suarez, H. (2004). "Long-term auditory adaptation to a modified peripheral frequency map," *Acta Oto-Laryngol.* **124**, 381–386.
- Sweller, J. (1994). "Cognitive load theory, learning difficulty, and instructional design," *Learn. Instr.* **4**, 295–312.
- Sweller, J. (1988). "Cognitive load during problem solving: Effects on learning," *Cogn. Sci.* **12**, 257–285.
- Van Gerven, P. W. M., Paas, F. G. W. C., Van Merriënboer, J. J. G., and Schmidt, H. G. (2002). "Cognitive load theory and aging: Effects of worked examples on training efficiency," *Learn. Instr.* **12**, 87–105.
- Verhaeghen, P., and Salthouse, T. A. (1997). "Meta-Analyses of Age-Cognition relations in adulthood: Estimates of linear and nonlinear age effects of structural models," *Psychol. Bull.* **122**, 231–249.
- Vermeire, K., Brokx, J. P., Wuyts, F. L., Cochet, E., Hofkens, A., and Van de Heyning, P. H. (2005). "Quality of life benefit from cochlear implantation in the elderly," *Otol. Neurotol.* **26**, 188–195.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (WAIS-3®)*, 3rd ed. (Harcourt Assessment, San Antonio TX).
- Wingfield, A., and Tun, P. (2007). "Cognitive supports and cognitive constraints on comprehension of spoken language," *J. Am. Acad. Audiol* **18**, 548–558.
- Xu, L., Thompson, C. S., and Pfingst, B. E. (2005). "Relative contributions of spectral and temporal cues for phoneme recognition," *J. Acoust. Soc. Am.* **117**, 3255–3267.