Speech recognition in normal hearing and sensorineural hearing loss as a function of the number of spectral channels

Deniz Başkent[a)]
House Ear Institute, Department of Auditory Implants, 2100 West Third Street, Los Angeles, California 90057

(Received 5 October 2005; revised 13 August 2006; accepted 14 August 2006)

Speech recognition by normal-hearing listeners improves as a function of the number of spectral channels when tested with a noiseband vocoder simulating cochlear implant signal processing. Speech recognition by the best cochlear implant users, however, saturates around eight channels and does not improve when more electrodes are activated, presumably due to reduced frequency selectivity caused by channel interactions. Listeners with sensorineural hearing loss may also have reduced frequency selectivity due to cochlear damage and the resulting reduction in the nonlinear cochlear mechanisms. The present study investigates whether such a limitation in spectral information transmission would be observed with hearing-impaired listeners, similar to implant users. To test the hypothesis, hearing-impaired subjects were selected from a population of patients with moderate hearing loss of cochlear origin, where the frequency selectivity would be expected to be poorer compared to normal hearing. Hearing-impaired subjects were tested for vowel and consonant recognition in steady-state background noise of varying levels using a noiseband vocoder and as a function of the number of spectral channels. For comparison, normal-hearing subjects were tested with the same stimuli at different presentation levels. In quiet and low background noise, performance by normal-hearing and hearing-impaired subjects was similar. In higher background noise, performance by hearing-impaired subjects saturated around eight channels, while performance by normal-hearing subjects continued to increase up to 12–16 channels with vowels, and 10–12 channels with consonants. A similar trend was observed for most of the presentation levels at which the normal-hearing subjects were tested. Therefore, it is unlikely that the effects observed with hearing-impaired subjects were due to insufficient audibility or high presentation levels. Consequently, the results with hearing-impaired subjects were similar to previous results obtained with implant users, but only for background noise conditions. [DOI: 10.1121/1.2354017]

PACS number(s): 43.71.Ky, 43.71.Rt, 43.66.Ts [KWG] Pages: 2908–2925

I. INTRODUCTION

Two factors that affect speech recognition by listeners with sensorineural hearing loss (SNHL) considerably are reduced audibility, due to elevated hearing thresholds, and suprathreshold deficits. In his model for recognition of speech in noise by hearing-impaired (HI) listeners, Plomp (1978) termed these factors as "attenuation" and "distortion," respectively. The attenuation problem can mainly be solved by amplification of sound to levels above the auditory thresholds. The solutions to the distortion problem are, however, not as straightforward. One form of distortion is related to reduced frequency selectivity, commonly seen in moderate to profound hearing loss of cochlear origin. The main cause for reduced frequency selectivity is thought to be damage in outer hair cells (OHCs). OHCs are believed to be an essential part of the active mechanism in cochlea that gives rise to cochlear nonlinearities such as the basilar membrane compression (Oxenham and Plack, 1997). The cochlear nonlinearities contribute to sharp tuning of the auditory filters (AFs) in normal hearing (NH). When there is damage in the cochlea, particularly in the OHCs, the nonlinearities might be reduced and the AFs might be broadened (Glasberg and Moore, 1986; Baker and Rosen, 2002).

Festen and Plomp (1983) observed that speech recognition by HI listeners in quiet was mainly affected by audiometric thresholds, while speech recognition in noise was affected by frequency resolution. The correlation between reduced frequency resolution and poor speech intelligibility in noise was later confirmed by other studies (Stelmachowicz *et al.*, 1985; Noordhoek *et al.*, 2000). In a study by Leek and Summers (1996), listeners with broad AFs were observed to need higher spectral contrast for discrimination of vowel-like stimuli presented in noise. It was hypothesized that a reduced SNR in the internal auditory representation of speech might be an explanation for the detrimental effects of reduced frequency resolution. Some studies also showed a correlation in quiet listening conditions; Turner and Henn (1989) observed that input filter patterns, a measure of frequency resolution, were correlated with recognition of vowels, and Henry *et al.* (2005) observed that the capability of resolving spectral peaks of rippled noise was correlated with recognition of vowels and consonants.

[a)]Present address: Starkey Hearing Research Center, 2150 Shattuck Avenue, Suite 408, Berkeley, California 94704. Electronic mail: deniz_baskent@starkey.com

Reduced frequency selectivity of hearing impairment was simulated with NH listeners as well. When speech was degraded by spectral smearing to simulate broadened AFs, speech recognition by NH listeners dropped, and the presence of the background noise amplified the detrimental effects (ter Keurs *et al.*, 1992, 1993; Baer and Moore, 1993; Boothroyd *et al.*, 1996).

Another method widely used to systematically explore the effects of temporal and spectral degradations on perception of speech is the noiseband vocoder (Shannon *et al.*, 1995; Xu *et al.*, 2005). In vocoder processing, speech signal is filtered into a number of spectral bands and narrow bands of noise are modulated with the envelopes extracted from individual bands. The final product is a synthesis of modulated noise bands. The processed speech has only crude spectral and temporal elements of the input speech. Cochlear implants (CIs) similarly deliver mainly speech envelope information, by modulating the current that stimulates the auditory nerve, and the gross spectral information, by stimulating distinct tonotopic places along the cochlea. Due to the similarity in the processing, despite the unknown percept of electrical stimulation, the noiseband vocoder has also been used to simulate CI processing with NH listeners.

Using a noiseband vocoder, spectral resolution of speech stimulus can be manipulated by changing the number of spectral channels. When vocoder speech was presented to NH subjects, performance increased as a function of the number of spectral channels. When CI users were tested under similar conditions, however, their performance usually saturated by 4–8 electrodes, and did not improve as more electrodes were activated (Fishman *et al.*, 1997; Friesen *et al.*, 2001; Garnham *et al.*, 2002). This finding suggests that, similar to HI listeners, implant users also have reduced frequency selectivity; they were unable to make use of the fine spectral resolution provided by the larger number of electrodes. However, unlike HI listeners, the main limiting factor for frequency resolution by implant users is believed to be channel interactions, caused by the summation of electric current fields or stimulation of the same nerve population by electrical pulses sent from different electrodes (Shannon, 1983).

Turner *et al.* (1995, 1999) used vocoder processing to explore the effects of reduced temporal or spectral resolution due to hearing impairment on perception of consonants in quiet. In the earlier study (Turner *et al.*, 1995), when a single-channel noiseband vocoder was used to eliminate all spectral cues, HI and NH listeners performed similarly. It was concluded that HI listeners were able to receive temporal information similar to NH listeners. In the second study, Turner *et al.* (1999) varied the number of channels from one to eight. The hypothesis was if the reduced frequency resolution of HI listeners was a limiting factor, the HI performance would reach an asymptote at a smaller number of channels compared to NH listeners, as it was the case with CI users. Performance by both subject groups continued to increase as a function of the number of spectral channels. However, except for the one-channel condition, performance by HI listeners was lower than NH listeners. Turner *et al.* concluded that even though HI listeners were able to make use of the temporal cues with one-channel processing, they were unable to combine the information from multiple channels as efficiently as NH listeners.

In the present study, it is hypothesized that the limitation can be observed if the experimental conditions of Turner *et al.* (1999) were expanded. There are several factors that might have prevented Turner *et al.* (1999) from fully observing the limiting effects of reduced spectral resolution on the speech recognition performance of HI listeners. The present study was designed by improving several factors: (1) In the Turner *et al.* (1999) study, the number of channels was increased only up to eight channels. Friesen *et al.* (2001) observed most difference in the performance by NH and CI subjects for number of channels higher than seven. For smaller number of channels, performance by implant users was lower than the performance by NH listeners, except for the single-channel condition, similar to findings by Turner *et al.* (1999). In the present study, the number of channels was varied from 2 to 40. (2) Previous studies showed strongest effects of reduced frequency selectivity due to hearing impairment on the perception of speech in noise (Festen and Plomp, 1983; Horst, 1987; Leek and Summers, 1996). Therefore, background noise of varying levels was added to the experimental conditions. (3) Perception of vowels has generally been observed to be more sensitive to spectral manipulations than consonants (Boothroyd *et al.*, 1996; ter Keurs *et al.* 1992; Turner and Henn, 1989). Therefore, the effects of changing the number of spectral channels could be stronger with vowel perception. In the present study, vowels were used as stimuli as well as consonants. (4) As mentioned above, audibility is one of the main factors affecting speech perception performance by HI listeners (i.e., Plomp, 1978). If it is not carefully controlled for, it might be difficult to separate the effects of suprathreshold deficits from the effects of audibility. Turner *et al.* (1999) maximized the audibility for HI listeners and used high presentation levels with NH listeners. In the present study, the audibility was similarly maximized for HI listeners. The presentation levels for NH listeners, however, were set at different levels to identify possible audibility effects. (5) It is also important to select the appropriate level of hearing loss. In the study by Turner *et al.* (1999), the baseline consonant recognition scores with unprocessed stimuli were considerably low with two out of six HI subjects. One of the two subjects with low scores had a severe hearing loss at high frequencies. Therefore, this subject probably did not have access to spectral information higher than 2 kHz, a region important for perception of consonants. In the present study, the inclusion criteria comprised having relatively flat hearing loss, from 50 to 60 dB HL, for frequencies up to 6 kHz. OHC damage was further confirmed by otoacoustic emission (OAE) measurements. To reduce possible audibility effects at higher frequencies, vocoder processing was limited to frequencies lower than 6 kHz.

The second purpose of the present study was to compare the results with HI listeners to the results with CI users (Friesen *et al.*, 2001).

HI listeners have inherently reduced frequency selectivity in the auditory system while the frequency selectivity of

TABLE I. Information about the HI subjects.

Subject	Age	Cause of the hearing loss	Age at diagnosis of the hearing loss (years)	Hearing aid user	Hearing thresholds (dB HL) of the test ear at the audiometric frequencies (Hz)					
					250	500	1000	2000	4000	6000
S1	34	German Measles	Shortly after birth	N	55	60	55	60	60	70
S2	62	Presbycusis	54	Y	40	50	60	50	50	65
S3	52	Streptomycin	8	Y	30	40	50	55	60	55
S4	63	Presbycusis	58	Y	45	55	55	50	55	65

CI users is thought to be mainly limited by channel interactions of the device processing. Therefore, many researchers proposed ways to minimize channel interactions in order to increase the effective number of spectral channels. For example, the "Continuous Interleaved Sampling" strategy, which delivers current pulses interleaved in time, was developed after strong channel interactions were observed with simultaneous activation of multiple electrodes (Wilson *et al.*, 1991). Electrode designs were improved to minimize current spread (e.g., Gstoettner *et al.*, 2001). Also different modes of stimulation, such as bipolar (BP) or tripolar (TP), have been suggested as an alternative to monopolar (MP) mode (i.e., Bierer and Middlebrooks, 2002). Many improvements in signal processing, hardware design, or stimulation methods have been shown to result in more localized stimulation patterns in physiological or psychophysical experiments. However, a beneficial effect of reduced channel interactions on the perception of speech has not been clearly demonstrated.

Studies on channel interactions may produce different results depending on the experimental design. In animal experiments, for example, the current level is usually set at a fixed dB level above threshold for all conditions. Stimuli are not loudness-balanced as this would require extensive training, and the dynamic range cannot be measured as the animals are incapable of reporting the maximum acceptable loudness levels. In BP and TP modes, the return electrodes are closer to the active electrode. Therefore, at the same current level, MP configuration would produce a wider stimulation pattern than the BP and TP configurations (Bierer and Middlebrooks, 2002). A higher degree of channel interaction was also observed with the MP configuration (Bierer and Middlebrooks, 2004). With human subjects, on the other hand, the stimuli can be loudness-balanced and the dynamic range can be measured. Usually, a small current is sufficient to produce the same loudness in MP mode as the BP or TP modes, as a larger population of nerves is stimulated in the MP mode due to the far positioning of the return electrode. When Kwon and van den Honert (2006) normalized the data using loudness-balancing and dynamic range measurements, the channel interaction patterns were similar for MP and BP modes. This observation might provide an explanation for why different stimulation modes do not always produce different speech recognition performance (i.e., Pfingst *et al.*, 1997; 2001).

The results with other measures of channel interactions have been mixed as well. Cazals *et al.* (1990) have not found a correlation between forward-masking patterns and speech recognition performance by CI users. Chatterjee and Shannon (1998) observed that one subject with the lowest sentence recognition scores had masking patterns with greatest dependence on masker level and on probe delay. Throckmorton and Collins (1999) measured channel interactions using electrode discrimination and forward masking and measured speech intelligibility with a number of stimuli of varying complexity. The results from both methods were correlated with some, but not all, measures of speech recognition. Stickney *et al.* (2006) showed that high levels of electric field interactions were correlated with low speech intelligibility performance for a simultaneous speech processing strategy (Simultaneous Analog Stimulation), but not for a sequential speech processing strategy (Continuous Interleaved Sampling).

The absence of a robust correlation between channel interactions (as measured in psychophysical or physiological studies) and speech recognition may imply that there might be other factors limiting the frequency selectivity, presumably originating in the auditory system, in addition to channel interactions caused by implant processing. In the present study, subjects with moderate SNHL also served as a model for an auditory system with inherently reduced frequency selectivity. It was explored if reduced frequency resolution in the auditory system could affect the performance in a manner similar to reduced frequency resolution observed with CI users.

II. METHODS

A. Subjects

Four HI subjects, aged 34–63 years (with an average of 52.75 years), participated in the study. All HI subjects were reported to have hearing loss of cochlear origin by their clinicians. All subjects other than S1 had symmetrical hearing loss. In the nontest ear, S1 had a mild hearing loss for low- to mid-frequencies, and a moderate hearing loss for higher frequencies; audiometric thresholds were rising from 35 to 25 dB HL for frequencies less than 1 kHz, sloping down from 25 to 55 dB HL for frequencies up to 4 kHz, and flat at 55 dB HL at higher frequencies. Detailed information about the HI subjects can be found in Table I.

Each subject was tested on one ear only. Test ears met the following criteria: (1) Relatively flat audiogram: A prescription method would have allowed customized amplification; however, it would also alter the spectral shape of speech. To prevent such additional factors that might affect the results, HI subjects with relatively flat audiogram were selected and the speech was made audible at all frequencies by changing the overall presentation level only.

(2) Pure tone thresholds around 50–60 dB HL for the audiometric frequencies of 250 Hz to 6 kHz: Moore (1996) suggested that for hearing loss up to 45 dB HL audibility is the most important factor for perception of speech; for higher degrees of hearing loss, effects of suprathreshold deficits can also be observed (Carney and Nelson, 1983). van Tasell (1993) suggested that a hearing loss less than 60 dB HL is generally associated with OHC loss. For these reasons, HI subjects were selected from a population of patients with hearing loss varying between 50–60 dB HL. These levels were optimal for the present study; the potential audibility effects could be minimized by proper amplification, and the suprathreshold deficits, presumably due to OHC loss, could still be observed.

(3) Nonfunctioning outer hair cells: One of the main assumptions in the present study is that the cochlear damage is primarily in the OHCs. OAEs are generally associated with the active mechanism in the cochlea and thought to represent the OHC function (Norton, 1992). Therefore, damage in OHCs was confirmed with (the absence of) OAEs.

(4) No dead regions: If there is additional damage in the inner hair cell or auditory nerve, the speech intelligibility performance would further be affected, and in possibly different ways depending on the pathology. To rule out dead regions, the Threshold Equalizing Noise (TEN) test (Moore *et al.*, 2000) was used. There was no elevation in the masked pure-tone thresholds with the TEN test for three subjects. The test was not applicable for subject S1 as the subject found the masking noise levels of the test uncomfortably high.

As a control group, five NH subjects, aged 27–35 years (with an average of 29.00 years), were tested with the same experimental conditions. NH subjects had thresholds better than 20 dB HL at audiometric frequencies from 250 Hz to 8 kHz. The right ear was chosen as the test ear for all NH subjects.

All subjects were native speakers of American English. The average pure tone thresholds of the test ears for both subject groups are shown in Fig. 1.

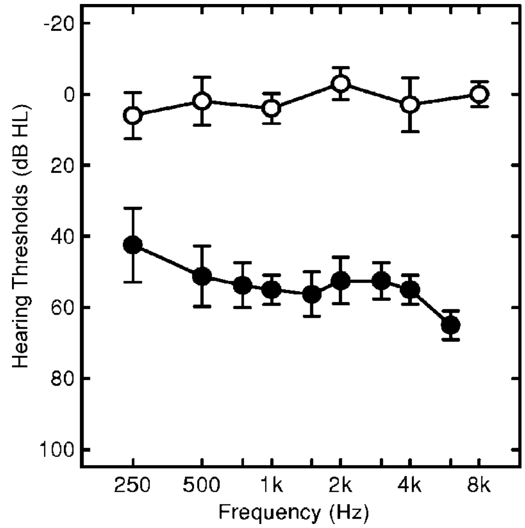

FIG. 1. Average pure tone hearing thresholds of the subjects, shown in dB HL. The open symbols show the thresholds of NH listeners for the audiometric frequencies from 250 Hz to 8 kHz, and the filled symbols show the thresholds of the HI listeners for the audiometric frequencies from 250 Hz to 6 kHz. Error bars show one standard deviation.

B. Stimuli

The speech recognition task was identification of medial vowels and consonants. Vowel stimuli (Hillenbrand *et al.*, 1995) consisted of twelve medial vowels presented in an /h/-vowel-/d/ context (heed, hid, head, had, hod, hawed, hood, who'd, hud, heard, hayed, hoed). The phonemes were spoken by 3 female and 3 male talkers and a total of 72 tokens were presented for each condition. Chance level on this test was 8.33% correct. Consonant stimuli (Shannon *et al.*, 1999) consisted of 20 medial consonants (/b tʃ d ð f g ʒ k l m n p r s ʃ t v w j z/), presented in an /a/-consonant-/a/ context. The phonemes were spoken by 2 female and 2 male talkers and a total of 80 tokens were presented for each condition. Chance level for this test was 5% correct.

For all stimuli, percent correct scores (PCS) were corrected for chance using the equation

$$PCS_{\text{corrected for chance}} = (PCS_{\text{raw}}\text{-Chance level})/(100\text{-Chance level}) \times 100.$$

In the corrected scores, a performance at 0% represents a performance at chance level.

In all experiments, a broadband noise with speech-shaped spectrum (SSN) was used. SSN was generated by filtering white noise using a Butterworth lowpass filter with a rollover of −12 dB\octave and a cutoff frequency of 800 Hz. The same SSN was also presented to the contralateral ear at a level 25 dB lower than the stimulus presentation level, to mask possible acoustic leakage to the nontest ear. The main motivation was to prevent one subject (S1) with asymmetrical hearing loss from listening with the better ear.

A Bruel and Kjaer artificial ear was used for calibration of the experimental setup. A reference tone signal at 1 kHz was generated with a total rms value of 65 dBA at the headphone output. The reference tone was used in setting the presentation level of stimuli to correct dBA values. The setting of the Crown Amplifier was fixed for the maximum presentation level of 95 dBA and a Tucker-Davis attenuator was used to produce the desired presentation levels. For the conditions with background noise, the speech and noise levels were proportionally adjusted to have the appropriate SNR, by using total rms values. The level of the combined signal was set to the presentation level. As a result, speech level decreased slightly as the SNR decreased.

C. Experimental procedure

The stimuli were processed with the noiseband vocoder (Shannon *et al.*, 1995) in real-time and presented using the CAST software created by Qian-Jie Fu at the House Ear Institute.

The overall frequency range of speech was limited to 200–6000 Hz. First, the stimulus was processed into a number of spectral channels using a set of Butterworth bandpass filters with a rollover of −24 dB\octave. The cutoff frequencies were determined by logarithmically dividing the input spectral range of 200–6000 Hz. The envelopes were extracted using half-wave rectification followed by a Butterworth lowpass filter with a cutoff frequency of 160 Hz and a rollover of −24 dB\octave. These bands were the analysis bands of the vocoder. The carrier noise bands were obtained by filtering wideband noise with the same set of bandpass filters. The envelopes from the analysis bands were used to modulate the carrier noise bands. In the last stage, modulated noise bands were combined to produce the processed stimulus, which at this point had only the coarse spectral and temporal information of the original stimulus and most fine structure was absent.

Phoneme recognition by NH and HI listeners was measured with unprocessed original stimuli and as a function of the number of spectral bands (2, 4, 6, 8, 10, 12, 16, 24, and 40 channels) of the vocoder, at varying background noise levels (quiet, SNR=10 dB, SNR=0 dB for vowels and consonants, and SNR=−5 dB for vowels only). The presentation level for the HI group was chosen to provide maximum audibility and comfort. NH listeners were tested at three presentation levels (experiments 1–3). The selection of the presentation levels is explained in detail in next section.

In experiment 4, an attempt was made to simulate the broad AFs with NH subjects. Dreschler and Plomp (1980) suggested two possible types of deterioration in the AF shape in SNHL: (1) critical bandwidth is broader; (2) slopes of the tuning curves are shallower. In experiment 2, the second type of deterioration was simulated by changing the filter slopes of the vocoder bandpass filters from −24 dB\octave to −6 dB\octave, while the cutoff frequencies remained the same. Shallower filter skirts resulted in considerable overlap between adjacent bands and produced spectrally smeared stimuli.

The speech stimuli were presented monaurally over Sennheiser HDA 200 headphones in a double-wall soundproof booth. A menu with the list of all possible phonemes was shown on the screen during testing. The subject identified the phoneme that was presented by selecting the appropriate entry in the menu using the mouse. The presentation order of individual tokens in each condition and the testing order of different conditions were randomized to minimize learning effects. As an additional caution, HI subjects were given one practice session of 2 h with similar testing procedure, where feedback was also provided, prior to actual data collection. NH subjects were familiar with noiseband vocoder technique from previous studies and therefore were not given a practice session. All subjects were allowed a preview of the stimuli at the beginning of each test.

NH subjects were tested once with each condition. Most HI subjects were tested more than once depending on subject's availability. Subjects S1 and S4 were tested twice for each condition. Subject S2 was tested three times with consonants for the noise conditions of 10 dB and 0 dB SNR, and three times with vowels for the noise conditions of 0 dB and −5 dB SNR.

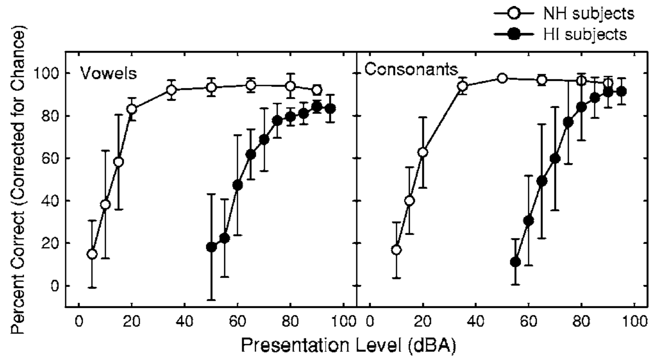

FIG. 2. Performance intensity (PI) functions. The average percent correct scores are plotted as a function of the stimulus presentation level. Open and filled symbols show the PI functions for NH and HI subjects, respectively. The stimuli were vowels and consonants, as shown in the left and right panels, respectively. The error bars show one standard deviation.

III. AUDIBILITY AND PRESENTATION LEVELS

The comfortable level for conversational speech is around 60–70 dB SPL. However, if speech was presented at an overall presentation level of 65 dB SPL in the present study, HI listeners would have insufficient audibility due to their elevated hearing thresholds. The presentation level for HI users, therefore, had to be set at higher levels to maximize audibility. On the other hand, the highest level could not exceed 100 dB SPL or so, as the stimuli would become uncomfortably loud due to loudness recruitment. For a fair comparison, the comfortable level of 65 dB SPL was not used with NH listeners either. Instead, the presentation levels were selected to be similar, either in sensation or in absolute levels, to the levels used with HI subjects.

The present study controls for the level effects using performance intensity (PI) functions, which were obtained by measuring percent correct scores with the test stimuli as a function of the overall presentation level. This method was preferred over using an amplification based on pure-tone thresholds or Speech Intelligibility Index (SII, ANSI S3.5-1997), as the PI functions provided a functional measure of how well the subjects performed with the specific stimuli used in the present study. The PI functions measured with unprocessed vowels and consonants are shown in the left and right panels, respectively, in Fig. 2. The open and filled symbols show the percent scores, corrected for chance, for NH and HI listeners, respectively.

The figure shows that NH listeners had a wide dynamic range for optimal presentation levels. From 35–40 dBA to 90 dBA, speech intelligibility was best and the loudness was reported to be comfortable. 50% of the peak PI levels (PI-50) was observed around 10–20 dBA. In contrast, speech intelligibility by HI listeners was best for high presentation levels, ranging from 80 to 95 dBA. For higher levels the stimulus was uncomfortably loud. PI-50 was observed for presentation levels around 60–65 dBA. As a result, the optimal listening levels were limited to the range from 80 to 95 dBA for HI subjects, significantly smaller than the range of optimal levels for NH listeners.

Note that even at the loudest presentation levels, the vowel recognition by HI listeners was slightly lower than the

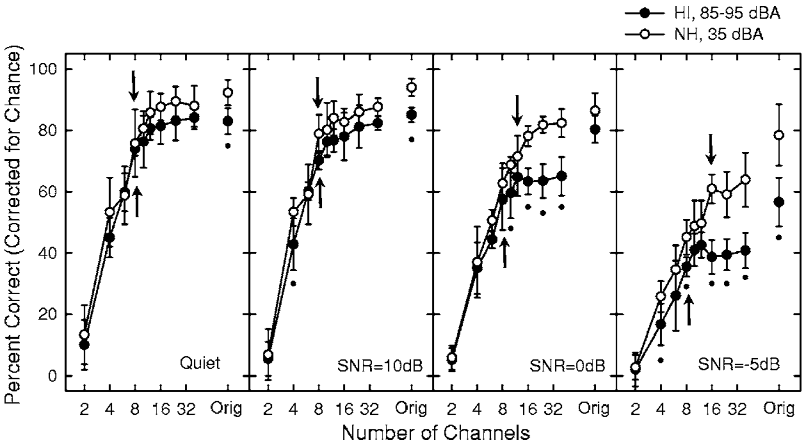

FIG. 3. Average vowel recognition scores, corrected for chance, shown as a function of the number of spectral channels. From left to right panel the background noise changes from quiet to SNR=−5 dB. The open and filled symbols show the performance by NH and HI listeners, respectively. The stars under the scores show the significant difference in the performance ($p<0.05$) by NH and HI listeners based on a posthoc Tukey multiple comparisons test, following a two-way mixed ANOVA. The performance by NH and HI groups for the unprocessed stimuli was compared using a *t*-test ($p<0.05$). The arrows above and below the data indicate the number of channels where the performance by NH and HI listeners, respectively, asymptoted.

NH listeners. When the vowel confusion matrices were analyzed, it was observed that all HI subjects and two out of five NH subjects had difficulty distinguishing the phoneme pair of "hod/hawed." Turner and Henn (1989) showed that recognition of vowels by HI listeners could be low if the formants of the vowels are similar. As the vowels used in their study were synthetically produced, the duration, which can be a robust cue for differentiating vowels with similar formants, was the same for all phonemes. For the pair of "hod/hawed," the formants are similar, and the duration cue is also not useful as both phonemes are of long duration. Differentiating these phonemes might have been a more difficult task for HI listeners than the NH listeners. The confusion caused by this one vowel pair could be an explanation for the slight difference observed in the performance by NH and HI listeners.

In the experiments, the presentation level for HI subjects was limited to the range from 85 to 95 dBA where all subjects had the best speech recognition. Subjects were given the flexibility to select the most comfortable level within this range. In general, higher levels were preferred for consonants than vowels. Subject S1, who had the poorest pure tone thresholds among all subjects, preferred louder levels.

NH subjects were tested at three presentation levels selected by comparing the PI functions of the two subject groups:

(A) Experiments 1 and 4: Comparable to HI subjects in sensation level (SL) with respect to PI-50; 35 dBA for vowels and 50 dBA for consonants.

The presentation levels of 85–95 dBA, at which the HI subjects were tested, were 25–30 dB higher than the presentation level at PI-50. For NH subjects, PI-50 was observed around 10 dBA for vowels, and around 20 dBA for consonants. The PI-50 presentation levels were defined as the reference value of 0 dB SL for each subject group. The presentation levels used for NH subjects were, then, made comparable to levels used with HI subjects in SL, by adding 25 dB to 10 dBA for vowels, and by adding 30 dB to 20 dBA for consonants.

(B) Experiment 2: Low level with reduced audibility; 20 dBA for vowels and 30 dBA for consonants.

To see the effects of decreased audibility the presentation levels were set to only 10 dB higher than the levels at PI-50. As a result, the levels were lower by 15–20 dB compared to experiment 1. It can be seen on the PI functions that the performance by NH subjects at these levels is poorer than the best performance (Fig. 2).

(C) Experiment 3: Comparable in absolute level, 85 dBA.

The shapes of AFs are level-dependent. As the level increases the filter shape becomes broader and more asymmetrical with an elevation on the low-frequency side (Glasberg and Moore, 2000). Therefore, even in NH subjects, frequency selectivity might be different at higher stimulus levels. A clear connection between the changes in AF shapes and overall speech perception has not been shown, yet several studies suggested that speech recognition by NH listeners might differ at high presentation levels. Studebaker *et al.* (1999), for example, observed a decrease in the intelligibility of speech in noise as the level increased from 64 to 99 dB SPL. Hornsby *et al.* (2005) observed a decrease in the transmission of consonant features at high levels. To account for potential effects of high presentation levels, NH listeners were tested at an absolute presentation level similar to the levels used with HI subjects.

IV. RESULTS

A. Experiment 1: Effect of the number of spectral channels

In experiment 1, presentation levels were 85–95 dBA for HI listeners. For NH listeners, the levels were set to 35 and 50 dBA, comparable to the levels selected for HI listeners in SL (as defined with respect to PI-50 levels), for vowel and consonant recognition tasks, respectively.

Percent correct scores for vowel recognition, averaged across subjects and corrected for chance, are plotted in Fig. 3, as a function of the number of spectral channels and for varying background noise levels. Performance by NH and HI groups is shown with open and filled symbols, respectively.

Performance by NH and HI listeners was compared using a two-may mixed Analysis of Variance (ANOVA), with the main factors of subject group and number of channels, and the interaction between the two factors. Performance by both groups increased as the number of the channels increased for all background noise conditions ($p<0.001$). The effect of the number of channels was similarly significant for all experiments reported in the present study, and therefore

TABLE II. F values of the two-way mixed ANOVA, used for comparing the performance by NH and HI subjects in experiments 1–3. The main effect of number of channels, which was significant for all experiments ($p < 0.001$), was not included. The main effect of subject group and the interaction between group and number of channels are shown for recognition of vowels and consonants, presented in varying background noise levels.

	Vowel recognition			Consonant recognition		
	Noise	Group factor $F(1,7)$	Group-channel interaction $F(8,56)$	Noise	Group factor $F(1,7)$	Group-channel interaction $F(8,56)$
Expt 1: Comfortable level	Quiet	3.10	0.37	Quiet	24.88[b]	0.83
	SNR=10 dB	3.94	1.03	SNR=10 dB	18.30[b]	7.73[b]
	SNR=0 dB	16.11[b]	3.15[b]	SNR=0 dB	17.40[b]	8.02[b]
	SNR=−5 dB	22.33[b]	4.13[b]			
Expt 2: Low level for NH listeners	SNR=0 dB	0.02	0.75	SNR=10 dB	0.74	2.22[a]
	SNR=−5 dB	2.23	1.87	SNR=0 dB	8.75[a]	3.59[b]
Expt 3: High level for NH listeners	SNR=0 dB	1.16	1.31	SNR=10 dB	41.68[b]	5.00[b]
	SNR=−5 dB	6.73[a]	2.66[a]	SNR=0 dB	55.30[b]	3.13[b]

[a] $p < 0.05$.
[b] $p < 0.01$.

will not be mentioned in reporting the results of the following experiments for simplicity. The F values are shown in Table II for the main factor of subject group and the interaction between group and number of channels. The F values with ‘a’ and ‘b’ denote significance at the levels of $p < 0.05$ and $p < 0.01$, respectively.

Following the ANOVA, a posthoc Tukey multiple comparisons test was used to compare all scores in pairs. The Tukey comparisons for the main factor of subject group within the same number of channels were used to identify the specific conditions where the performance by NH and HI listeners differed significantly. The scores with the unprocessed stimuli were not included in the ANOVA, therefore they were compared using a t-test. In Fig. 3, the conditions where the difference in the scores was statistically significant ($p < 0.05$) are shown by the dots under the scores. The Tukey comparisons for the main factor of number of channels within NH or HI subjects were used to identify the number of channels where the performance by that subject group reached the asymptote. For each subject group, the highest percent correct scores were used in the comparisons. The lowest number of channels where the performance did not differ significantly compared to the highest percent correct scores was accepted as the saturation point. In Fig. 3, the number of channels where the performance reached the asymptote is shown by arrows above the data, for NH listeners, and below the data, for HI listeners. Table III summarizes the number of channels at the saturation point for all experiments of the present study.

There was no significant main effect of group on vowel recognition in quiet and at the low noise level of SNR =10 dB. For both NH and HI subjects, performance reached the asymptote around 8 channels (Fig. 3, left panels; Table III). However, performance by NH and HI subjects differed significantly when the background noise level was higher; there was a significant effect of subject group ($p < 0.01$) and a significant interaction ($p < 0.01$) between group and number of channels, as shown in Table II for SNR=0 dB and SNR=−5 dB. As the noise level increased, NH listeners were able to employ a higher number of spectral channels. NH performance increased up to 12 and 16 channels for SNR=0 dB and SNR=−5 dB, respectively, while performance by HI listeners saturated at 8 channels for all noise conditions (Fig. 3, right panels; Table III).

Figure 4 shows the average percent correct scores for consonant recognition, corrected for chance, as a function of the number of spectral channels and for varying background noise levels.

There was a significant main effect of subject group on

TABLE III. The summary for the number of channels where vowel and consonant recognition by NH and HI subjects saturated.

	Vowels				Consonants		
	Quiet	SNR= 10 dB	SNR= 0 dB	SNR= −5 dB	Quiet	SNR= 10 dB	SNR= 0 dB
HI Expt 1: Comfortable level	8	8	8	8	8	8	8
NH Expt 1: Comfortable level	8	8	12	16	8	10	12
NH Expt 2: Low level	–	–	12	12	–	10	16
NH Expt 3: High level	–	–	12	12	–	12	8
NH Expt 4: Wide carrier filters	–	–	16	–	–	16	–
NH Expt 4: Wide carrier and analysis filters	–	–	24	–	–	24	–

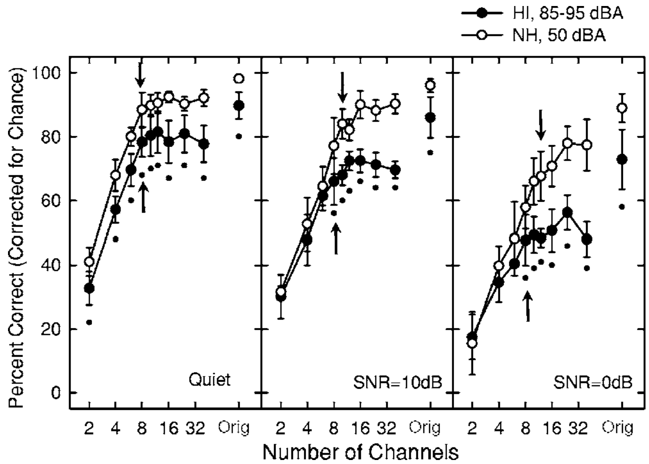

FIG. 4. Similar to Fig. 3, except for the stimuli were consonants and the background noise increases from quiet to SNR=0 dB, from left to right panels. The open and filled symbols show the scores by NH and HI subjects, respectively. The dots under the scores denote the specific conditions where the performance was significantly different ($p<0.05$) between NH and HI listeners. The arrows show the number of channels where the performance by a specific subject group reached the asymptote.

consonant recognition in all background noise conditions ($p<0.01$, as shown in Table II). In quiet, however, the patterns were similar as the interaction between group and channels was not significant, and performance by both groups reached the asymptote around 8 channels (Fig. 4, left panel; Table III). At higher noise levels, on the other hand, the patterns were significantly different; there was a significant interaction of group and channels at SNR=10 dB and SNR =0 dB ($p<0.01$, as shown in Table II). As the background noise increased from SNR=10 dB to SNR=0 dB, performance by NH subjects increased up to 10 and 12 channels, respectively, while performance by HI listeners saturated by 8 channels for both noise conditions (Fig. 4, right panels; Table III).

Note that, for both vowels and consonants, there was a difference in the performance by NH and HI listeners even in the quiet and unprocessed conditions. The difference in the performance with unprocessed stimuli generally increased as the noise level increased; the drop in the performance by HI listeners was sharper than the drop in the performance by NH listeners as a function of the background noise level. When the noiseband vocoder processing was added, there was generally a further drop from the performance with the unprocessed condition, even at the highest number of channels used in the present experiment. The difference in the performance between the 40-channel and unprocessed conditions are shown in percent correct scores in Table IV, for both subject groups and for both vowels and consonants. A paired t-test was used to determine statistical significance. The drop in the performance from the unprocessed condition to the 40-channel processing condition also increased with increasing noise, and this effect was generally more pronounced with HI listeners (Table IV).

B. Experiment 2: Effects of audibility

In experiment 2, the effect of reduced audibility on recognition of vocoder speech was explored. The presentation level for NH listeners was reduced by 15–20 dB compared to experiment 1; the vowels were presented at 20 dBA and the consonants were presented at 30 dBA. Only the noise conditions from experiment 1 where the patterns of the performance by two subject groups were significantly different were repeated. The average percent correct scores at the low presentation levels are presented in Fig. 5, superimposed with the scores from experiment 1. Figure 5(A) shows the results for vowel recognition with background noise of SNR=0 dB and SNR=−5 dB in the left and right panels, respectively. Figure 5(B) shows the scores for consonant recognition with SNR=10 dB and SNR=0 dB in the left and right panels, respectively. In each panel, open and filled circles replicate the scores by NH and HI subjects, respectively, from experiment 1. Open triangles show the scores by NH subjects when the stimuli were presented at lower levels.

The performance by NH listeners at the low presentation level and the performance by HI listeners from experiment 1 were compared using a two-may mixed ANOVA, with the main factors of subject group and number of channels, and the interaction between the two factors. The F values are shown in Table II for the main factor of group and the interaction between group and number of channels. A Tukey test was used to compare the scores with the vocoder processed conditions, and a t-test was used to compare the scores with the unprocessed conditions. The dots under the data show the conditions where the performance by NH listeners at the low presentation level and the performance by HI subjects differed significantly ($p<0.05$). The Tukey test was also used to determine the number of channels where the performance reached the asymptote, as it was shown by arrows above the data, for NH listeners, and below the data, for HI listeners (Fig. 5).

The effect of presentation level on performance by NH listeners was explored by comparing the scores measured at

TABLE IV. The difference in performance between unprocessed and 40-channel processing conditions shown in percent correct score for vowels and consonants and for each subject group.

	Vowels				Consonants		
	Quiet	SNR= 10 dB	SNR= 0 dB	SNR= −5 dB	Quiet	SNR= 10 dB	SNR= 0 dB
NH subjects	4.40%	6.36%[a]	3.94%	14.54%[a]	6.05%[a]	5.79%[a]	11.58%
HI subjects	1.08%	2.65%[b]	15.26%[b]	15.82%	11.99%[b]	16.34%[b]	24.89%[b]

[a]Paired t-test: $p<0.05$.
[b]Paired t-test: $p<0.01$.

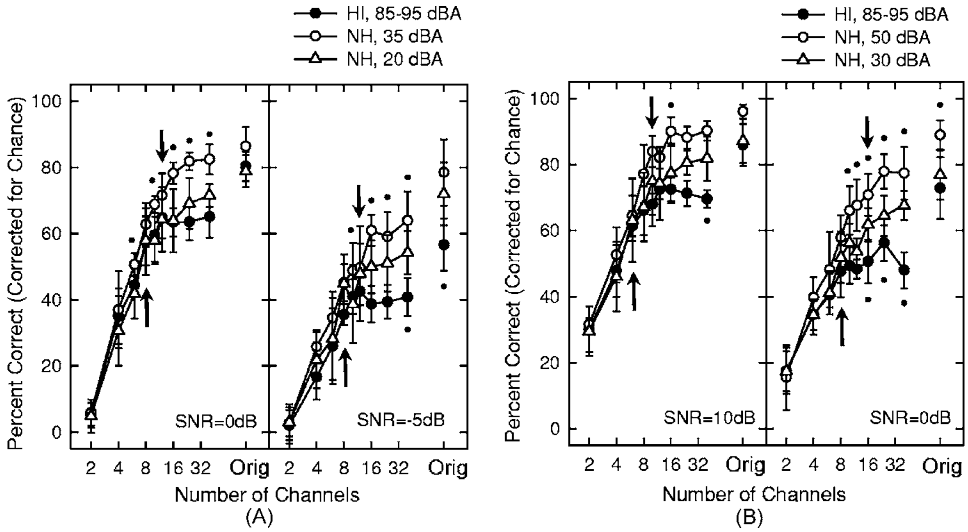

FIG. 5. Average phoneme recognition scores by NH listeners from experiment 2, where the presentation level was low, presented with the results from experiment 1, where the presentation level was comfortable. (A) shows the vowel recognition scores by HI listeners (filled circles, replicated from Fig. 3) and by NH listeners with stimuli presented at 35 dBA (open circles, replicated from Fig. 3) and at the low level of 20 dBA (open triangles). Left and right panels show the results with background noise at SNR=0 dB and SNR=−5 dB, respectively. (B) shows the consonant recognition scores by HI listeners (filled circles, replicated from Fig. 4) and by NH listeners when the stimuli were presented at 50 dBA (open circles, replicated from Fig. 4) and at the low level of 30 dBA (open triangles). Left and right panels show the results with background noise at SNR=10 dB and SNR=0 dB, respectively, The small dots above the open circles indicate the conditions where the performance by NH subjects dropped significantly as a result of reduced presentation level, shown with a Tukey test ($p < 0.05$). The small dots under the data indicate the conditions where the performance by NH subjects at the reduced presentation level was significantly different than the performance by HI subjects, shown with a Tukey test ($p < 0.05$). The conditions with unprocessed original stimuli were compared with a *t*-test ($p < 0.05$). The arrows above and below the data indicate the number of channels where the performance by NH and HI listeners, respectively, asymptoted.

the low presentation level of experiment 2 and the scores measured at the comfortable level of experiment 1. A two-way repeated-measures (RM) ANOVA with the main factors of presentation level and number of channels was used. The corresponding F values are presented for the main factor of presentation level and the interaction between level and number of channels in Table V. The dots above the data show the conditions where the performance by NH listeners at two presentation levels differed significantly, as shown by the Tukey or *t*-test ($p < 0.05$).

The results, presented in Fig. 5, show that reduced audibility generally produced lower scores than the scores at comfortable level. In unprocessed conditions, performance by NH listeners at low presentation level was more similar to performance by HI listeners. Even with the highest number of channels, performance with processed stimuli was lower than the unprocessed conditions for both subjects groups; however, the drop in performance with vocoder processing was generally larger for HI listeners.

In the conditions with processed vowels, there was a significant main effect of presentation level on NH performance at SNR=0 dB noise level ($p < 0.05$, Table V). No significant effect of level at SNR=−5 dB, and no significant interaction at both background noise levels were observed. The lower scores by NH listeners, obtained in experiment 2 as a result of the low presentation level of 20 dBA, were not

TABLE V. F values of the two-way RM ANOVA. Performance by NH listeners at low and high presentation levels (experiments 2 and 3, respectively) was compared to performance by NH listeners at comfortable levels (experiment 1). The main effect of the presentation level and the interaction between level and number of channels are shown for recognition of vowels and consonants, presented in varying background noise levels.

	Vowel recognition			Consonant recognition		
	Noise	Level factor $F(1,4)$	Level-channel interaction $F(8,32)$	Noise	Level factor $F(1,4)$	Level-channel interaction $F(8,32)$
Expt 2: Low level for NH listeners	SNR=0 dB	15.53[a]	1.33	SNR=10 dB	2.97	1.50
	SNR=−5 dB	7.30	1.63	SNR=0 dB	10.83[a]	2.59[a]
Exp 3: High level for NH listeners	SNR=0 dB	4.12	1.23	SNR=10 dB	0.66	2.39[a]
	SNR=−5 dB	1.67	1.65	SNR=0 dB	2.77	2.41[a]

[a]$p < 0.05$.

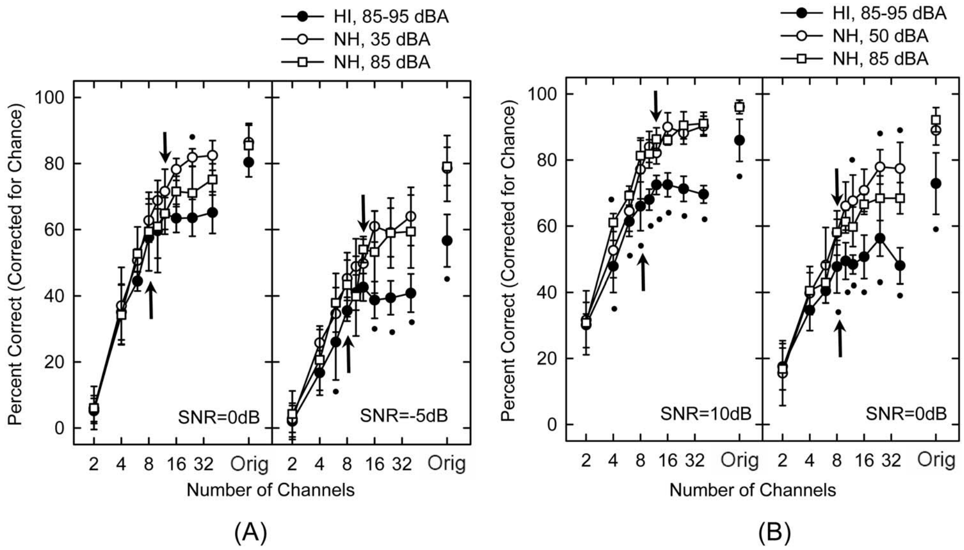

FIG. 6. Similar to Fig. 5, the filled and open circles show the scores by HI and NH listeners, respectively, replicated from experiment 1. The open squares show the percent correct scores by NH listeners with stimuli presented at the high level of 85 dBA.

significantly different than the scores by HI listeners (Table II). However, performance by NH subjects at this low level increased up to 12 channels; a larger number than 8 channels where performance by HI listeners reached the asymptote [Fig. 5(A); Table III].

In the conditions with processed consonants, the low presentation level did not change the performance by NH listeners at SNR=10 dB significantly; however, it resulted in a significant drop in scores at SNR=0 dB compared to the comfortable level, as shown by the significant main effect of level and the interaction ($p<0.05$, Table V). Performance by NH listeners at the low level of 30 dBA was similar to performance by HI listeners for the SNR=10 dB noise level; but at SNR=0 dB there was a significant main effect of subject group ($p<0.05$, Table II). The interaction between group and channels was significantly different at both noise levels ($p<0.05$ at SNR=10 dB and $p<0.01$ at SNR=0 dB, Table II), indicating a difference in the trends of the scores by NH and HI listeners. The Tukey test further showed that performance by NH subjects at the low presentation level increased up to 10 and 16 channels for the background noise conditions of SNR=10 dB and SNR=0 dB, respectively, while performance of HI subjects reached the asymptote at 8 channels [Fig. 5(B); Table III].

C. Experiment 3: Effects of high presentation level

In experiment 3, NH subjects were tested at a high presentation level of 85 dBA, comparable to the high presentation levels of 85–95 dBA used with HI listeners. The average percent correct scores at high levels, corrected for chance, are shown in Fig. 6 presented with scores from experiment 1. The open squares show the scores by NH listeners at the high presentation level of 85 dBA. The open circles show the scores by NH listeners from experiment 1, where vowels and consonants were presented at the comfortable levels of 35 and 50 dBA, respectively. The filled circles show the scores by HI listeners from experiment 1. The vowel recognition scores at SNR=0 dB and SNR=−5 dB are presented in the left and right panels, respectively, of Fig. 6(A). The consonant recognition scores at SNR=10 dB and SNR=0 dB are presented in the left and right panels, respectively, of Fig. 6(B). The dots under the data show the conditions where the performance differed between NH and HI listeners, when both subject groups were tested at high presentation levels. The dots above the data show the conditions where the performance by NH listeners at comfortable and high presentation levels was different. The arrows below and above the data show the number of channels where the performance by HI and NH listeners at high levels, respectively, saturated.

In the unprocessed conditions, the scores by NH listeners did not change when the presentation level was increased from the comfortable levels of 35 and 50 dBA to the loud level of 85 dBA.

In the conditions with processed vowels, increasing the presentation level did not change the scores by NH listeners significantly; at both background noise conditions there was no significant main effect of level and no significant interaction (Table V). When performance at high presentation level was compared to performance by HI listeners, there was also no significant main effect of group and no interaction of group and channels at SNR=0 dB; however, performance by NH listeners saturated around 12 channels while performance by HI listeners saturated around 8 [Fig. 6(A), left panel; Table III]. There was a significant main effect of group and significant interaction at SNR=−5 dB ($p<0.05$, Table II). The number of channels at saturation point was 12 for NH listeners, and 8 for HI listeners [Fig. 6(A), right panel; Table III].

In the conditions with processed consonants, increasing the presentation level did not change the consonant recognition scores by NH listeners significantly; however, there was significant interaction between level and number of channels in both background noise levels ($p<0.05$, Table V). The scores by NH and HI listeners tested at similar absolute levels were significantly different in both background noise levels; there was a significant main effect of group and significant interaction ($p<0.01$, Table II). At SNR=10 dB, performance by NH listeners asymptoted at 12 channels while performance by HI listeners asymptoted at 8 channels

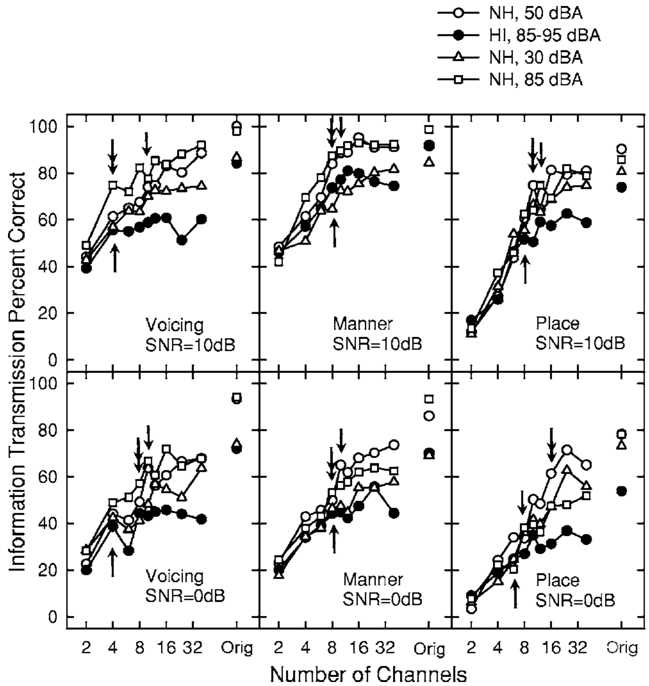

FIG. 7. Information transmission scores shown for the production based categories of voicing, manner, and place of articulation. The top and bottom rows present the scores obtained in the presence of background noise with SNR=10 dB and SNR=0 dB, respectively. The filled circles show the scores by HI subjects. The open symbols show the scores by NH subjects for three different presentation levels; open circles for 50 dBA, open triangles for 30 dBA, and open squares for 85 dBA. Arrows above the data show the number of channels at the asymptotic performance by NH listeners at different presentation levels. Arrows below the data show the number of channels at the asymptote for HI listeners.

[Fig. 6(B), left panel; Table III]. At SNR=0 dB, performance by both subject groups asymptoted at 8 channels [Fig. 6(B), right panel; Table III].

D. Analysis of consonant features

The ratio of spectral to temporal information required for the perception of the consonant features increases from voicing to manner to place of articulation (van Tasell *et al.*, 1992; Turner *et al.*, 1995). ter Keurs *et al.* (1992) and Boothroyd *et al.* (1996), for example, showed that spectral smearing affected the perception of the place feature most. To observe the effect of vocoder processing on perception of these features, consonant confusion matrices of experiments 1–3 were analyzed for the production based categories of voicing, manner, and place of articulation (Miller and Nicely, 1955).

The information transmission percent correct scores are presented for the noise conditions of SNR=10 dB and SNR=0 dB in the top and bottom panels of Fig. 7, respectively. The panels from left to right show the scores for voicing, manner, and place of articulation. Filled circles show the scores by HI subjects. Open symbols show the scores by NH subjects for three presentation levels; open circles for 50 dBA, open triangles for 30 dBA, and open squares for 85 dBA. The standard deviations were omitted for clarity.

The *F* values of the two-way mixed ANOVA that was used to compare the perception of features by NH and HI listeners are presented in Table VI. The *F* values of the two-way RM ANOVA that was used to compare the perception of features by NH listeners at different presentation levels are presented in Table VII. The number of channels where the performance reached the asymptote was determined by the Tukey test. These values are summarized in Table VIII, and are also shown in Fig. 7 by the arrows above and below the data for NH and HI listeners, respectively.

The left panels of Fig. 7 show that the voicing transmission patterns by NH listeners were similar at all presentation levels (Table VII) and performance by NH listeners was generally better than HI listeners (Table VI). The two-way mixed ANOVA showed that the interaction between group and channels was significantly different only in experiment 1 where the presentation level was set to the comfortable level of 50 dBA for NH listeners. The Tukey test confirmed this finding; performance by NH listeners reached the asymptote at 10 channels while performance by HI listeners reached the asymptote at 4 channels (Table VIII), in both background noise levels. In experiments 2 and 3, where the presentation levels for NH listeners were low (30 dBA) and high (85 dBA), respectively, both subject groups had the

TABLE VI. *F* values of the two-way mixed ANOVA, used to compare the performance by NH and HI subjects. The main effect of group and the interaction between group and channels are shown for production based categories of voicing, manner, and place of articulation for experiments 1–3.

		Voicing		Manner		Place	
	Noise	Group factor $F(1,7)$	Interaction $F(8,56)$	Group factor $F(1,7)$	Interaction $F(8,56)$	Group factor $F(1,7)$	Interaction $F(8,56)$
Expt 1: Comfortable level	SNR=10 dB	5.53[a]	2.75[a]	17.32[b]	2.06	46.59[b]	6.76[b]
	SNR=0 dB	3.58	2.34[a]	27.18[b]	4.98[b]	42.47[b]	7.36[b]
Expt 2: Low level	SNR=10 dB	2.25	0.84	0.21	1.74	2.13	1.77
	SNR=0 dB	9.62[a]	1.14	1.47	1.64	11.35[a]	4.09[b]
Expt 3: High level	SNR=10 dB	21.99[b]	1.89	59.39[b]	3.52[b]	20.34[b]	3.68[b]
	SNR=0 dB	27.90[b]	1.11	37.79[b]	1.90	12.13[b]	1.83

[a] $p<0.05$.
[b] $p<0.01$.

TABLE VII. F values of the two-way RM ANOVA, used to compare the performance by NH subjects at low and high presentation levels to the performance at comfortable level. The main effect of level and the interaction between level and channels are shown for production based categories of voicing, manner, and place of articulation for experiments 2 and 3.

		Voicing		Manner		Place	
	Noise	Level factor $F(1,4)$	Interaction $F(8,32)$	Level factor $F(1,4)$	Interaction $F(8,32)$	Level factor $F(1,4)$	Interaction $F(8,32)$
Expt 2:	SNR=10 dB	0.87	0.57	5.75	2.39[a]	0.32	1.38
Low level	SNR=0 dB	0.98	1.31	23.64[b]	1.95	11.48[a]	1.24
Expt 3:	SNR=10 dB	4.72	0.88	0.62	2.06	0.03	3.74[a]
High level	SNR=0 dB	0.93	0.57	3.00	1.62	12.23[a]	1.05[b]

[a]$p < 0.05$.
[b]$p < 0.01$.

asymptotic performance at 4 channels at SNR=10 dB noise level. At the higher noise level of SNR=0 dB, performance by NH subjects continued to increase up to 8–10 channels (Table VIII).

The middle panels of Fig. 7 show that the transmission of manner feature by NH listeners was similar at the levels of 50 and 85 dBA (Table VII), and at both levels NH listeners performed significantly better than HI listeners (Table VI). However, when the presentation level was reduced to 30 dBA, the performance dropped; the perception of manner by NH subjects at the low level of 30 dBA was similar to perception of manner by HI listeners (Table VI). The two-way mixed ANOVA showed a significant interaction of hearing loss and channels in experiment 1 for SNR=0 dB and experiment 3 for SNR=10 dB. However, the overall results of the Tukey test showed that the number of channels at performance saturation was similar (8 or 10) for both subject groups and in all experiments (Table VIII).

The right panels of Fig. 7 show that the transmission of place by NH listeners was similar at all presentation levels for SNR=10 dB. At SNR=0 dB, performance at the comfortable level of 50 dBA was significantly better than performance at the lower and higher presentation levels of 30 and 85 dBA (Table VII). There was also significant interaction between level and channels when the performance by NH listeners was compared at 50 and 85 dBA (Table VII). Performance by NH subjects was generally better than the HI subjects, and there was a significant interaction between group and channels in most settings (Table VI). At SNR =10 dB, the number of channels where the performance saturated was slightly higher for NH listeners (10–12) compared to HI listeners (8; as shown in Table VIII). At SNR =0 dB, perception of place cue by HI listeners increased up to 6 channels. Place perception by NH listeners, on the other hand, increased up to 16 channels at the presentation levels of 30 and 50 dBA, and up to 8 channels at the presentation level of 85 dBA (Table VIII).

E. Experiment 4: Simulation of broad auditory filters

In experiment 4, the broad AFs were simulated with NH subjects by changing the filter slopes of the vocoder filters from a rollover of −24 dB\octave to a rollover of −6 dB\octave. In the first part, only the carrier filters were changed while the same analysis bands were used; this setup can be interpreted as a simulation of an input signal with good spectral resolution while the receiving component of the transmission system has reduced spectral resolution. In the second part, both carrier and analysis filters were made wider. This setup can be interpreted as a simulation of a transmission system with overall reduced spectral resolution.

The average vowel and consonant recognition scores, corrected for chance, are shown in Figs. 8(A) and 8(B), respectively, as a function of the number of spectral channels.

TABLE VIII. The number of channels where the performance by NH and HI listeners are saturated, shown for the production based categories of voicing, manner, and place of articulation.

	Voicing		Manner		Place	
	SNR=10 dB	SNR=0 dB	SNR=10 dB	SNR=0 dB	SNR=10 dB	SNR=0 dB
HI Expt 1 Comfortable level	4	4	8	8	8	6
NH Expt 1 Comfortable level	10	10	10	10	10	16
NH Expt 2 Low level	4	10	10	8	10	16
NH Expt 3 High level	4	8	8	8	12	8

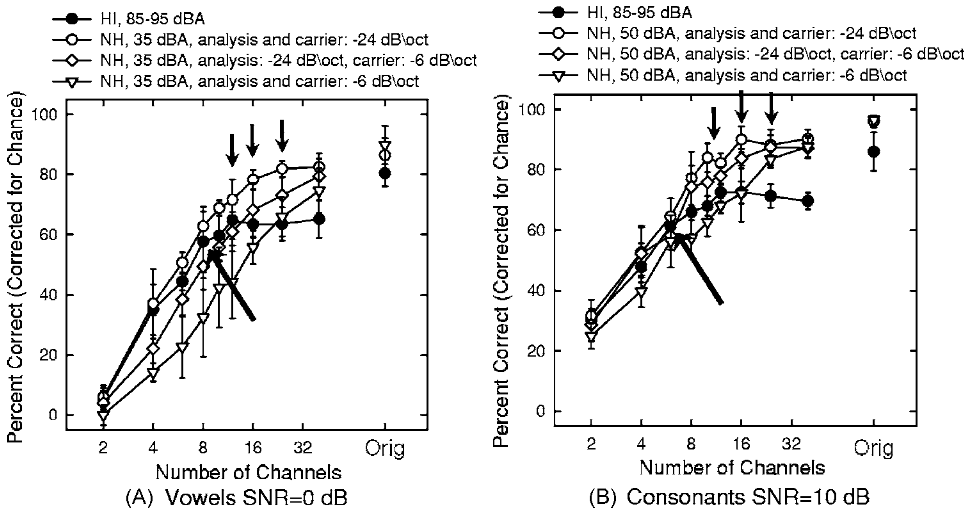

FIG. 8. Average phoneme recognition scores by NH listeners when tested with wide vocoder filters. The filled and open circles show the scores by HI and NH listeners, respectively, replicated from experiment 1. (A) and (B) show the vowel recognition scores with a background noise of SNR=0 dB and the consonant recognition scores with a background noise of SNR=10 dB, respectively. The open diamonds and the open triangles show the scores by NH listeners when the carrier filter slope was made shallower and when both the carrier and analysis filter slopes were made shallower, respectively. The arrows above the data show the number of channels where the performance by NH listeners reached the asymptote. The arrows below the data show the same for HI listeners.

The open and filled circles replicate the scores of HI and NH subjects, respectively, from experiment 1. The open diamonds show the average scores by NH subjects where only the carrier filter slopes were made shallower. The open triangles show the performance where both carrier and analysis filter slopes were made shallower. Vowels were presented in the background noise of SNR=0 dB, and consonants were presented in the background noise of SNR=10 dB. The presentation levels of the phonemes were the same comfortable level as in experiment 1; 35 dBA for vowels, and 50 dBA for consonants.

A two-way mixed ANOVA was used to compare the results by NH listeners tested with wide vocoder filters and the results by HI listeners. A two-way RM ANOVA was used to compare the results by NH listeners with narrow and wide filters. The corresponding *F* values are presented in Table IX. The number of channels at the saturation point is shown in Table III, and is denoted by arrows above and below the data for NH and HI listeners, respectively, in Fig. 8.

When only the carrier filters were made wider, the performance by NH listeners did not change significantly (as shown by the open diamonds in Fig. 8 and also in the lower portion of Table IX). The number of channels where the performance saturated, however, increased from 10–12 to 16 (Table III). A stronger effect was observed when carrier and analysis filters were manipulated together (as shown by open triangles in Fig. 8). The scores by NH listeners dropped significantly and there was a significant interaction between fil-

TABLE IX. The upper portion; the *F* values of the two-way mixed ANOVA used to compare the performance by NH listeners with wide vocoder filters, from experiment 4, with the performance by HI listeners, from experiment 1. The lower portion; the F values of the two-way RM ANOVA used to compare the performance by NH listeners with narrow vocoder filters, from experiment 1, and wide vocoder filters, from experiment 4.

	Vowel recognition SNR=0 dB		Consonant recognition SNR=10 dB	
Two-way mixed ANOVA	Group factor $F(1,7)$	Interaction $F(8,56)$	Group factor $F(1,7)$	Interaction $F(8,56)$
Wide carrier filters	0.08	5.47	8.13[a]	6.41[b]
Wide carrier and analysis filters	11.28[a]	7.43[b]	0.12	9.02[b]
	Vowel recognition SNR=0 dB		**Consonant recognition SNR=10 dB**	
Two-way RM ANOVA	Filter slope factor $F(1,4)$	Interaction $F(8,32)$	Filter slope factor $F(1,4)$	Interaction $F(8,32)$
Wide carrier filters	2.87	2.55[a]	2.04	0.84
Wide carrier and analysis filters	25.52[b]	6.68[b]	29.45[b]	5.54[b]

[a] $p<0.05$.
[b] $p<0.01$.

ter bandwidth and channels (as shown in the lower portion of Table IX). The number of channels where the performance saturated increased from 10–12 to 24 (Table III).

Despite the substantial changes observed in the performance by NH listeners, either method failed to reproduce the trends seen in data by HI listeners. There was an overall reduction in the scores and the difference between the scores by NH and HI listeners was smaller as a result of these manipulations. However, the interaction between group and channels was generally significant (Table IX), indicating that the trends in the data by the two subject groups were different. The number of channels at the saturation point was higher with NH listeners than with HI listeners; 16 or 24 compared to 8 channels (Table III).

V. DISCUSSION

Turner *et al.* (1999) measured consonant recognition in quiet by NH and HI listeners as a function of the number of vocoder channels. The number of the channels varied from one to eight and the performance by HI listeners was poorer than NH listeners for all conditions, except for the single-channel processing. Yet, the performance by both subject groups increased as a function of the number of channels, contrary to the expectation that the performance by HI listeners would saturate at a smaller number of channels due to the limiting effect of reduced spectral resolution.

Friesen *et al.* (2001), however, observed a limiting effect of reduced spectral resolution on the perception of speech by CI users. When recognition of phonemes and sentences, processed with a noiseband vocoder, was measured with NH listeners as a function of the number of spectral channels, the performance increased up to 20 channels. Under similar listening conditions, the performance by CI users saturated around seven electrodes.

In the present study, it was hypothesized that a similar limiting effect of reduced spectral resolution might also be observed with HI listeners if the conditions from Turner *et al.* (1999) study were extended to a greater number of channels. Vowels were included in the stimuli as they were hypothesized to be more sensitive to spectral manipulations. The number of channels varied from 2 to 40, an upper limit selected on purpose higher than eight channels where the saturation in performance was observed with best CI users. Similar to the Friesen *et al.* (2001) study, background noise was added, as many studies showed that speech recognition by HI listeners in noise might be correlated with spectral resolution. Additionally, strict inclusion criteria were followed in the recruitment of the HI subjects and audibility was controlled by testing NH listeners at varying presentation levels.

In experiment 1, phoneme recognition was measured at similar sensation levels, determined from performance intensity functions by taking PI-50 levels as the reference for each subject group. Vowel recognition by NH and HI listeners, measured as a function of the number of vocoder channels, were similar in quiet and at the low background noise level of SNR=10 dB; the performance by both subject groups increased up to 8 channels before reaching plateau. Results with consonant recognition in quiet were consistent with results by Turner *et al.* (1999); HI listeners generally had lower scores compared to NH listeners, but the trend in the performance was similar.

The limiting effects of reduced spectral resolution, similar to that reported by Friesen *et al.* (2001) with CI subjects, were observed with HI subjects only when the background noise level was increased. In all noise conditions, the performance by HI listeners saturated by 8 channels. Performance by NH listeners, on the other hand, increased up to 12 and 16 channels with vowel recognition at the noise levels of SNR =0 dB and SNR=−5 dB, respectively, and up to 10 and 12 channels with consonant recognition at SNR=10 dB and SNR=0 dB, respectively. This finding implies that while NH listeners could utilize a higher number of channels in challenging listening situations, such as when background noise was added, HI listeners did not show such ability.

The confusion matrices, obtained with consonants in background noise, were further analyzed for the transmission of the production based categories of voicing, manner, and place of articulation. Perception of the place feature is generally most sensitive to spectral degradations; therefore the effects were expected to be seen mainly with place. Consistent with this expectation, perception of place differed for NH and HI listeners; performance by HI listeners saturated by 6–8 channels while performance by NH listeners saturated by 10–16 channels. On the other hand, because perception of voicing and manner rely heavily on the temporal cues, the difference in the performance by NH and HI listeners was expected to be smaller with these features. Perception of manner was similar by the two subject groups. However, contrary to expectations, perception of voicing differed for NH and HI listeners; the number of channels at the saturation point was higher with NH listeners than with HI listeners when measured at similar SL (10 vs 4 channels).

A. Presentation levels

Audibility has a significant effect on speech recognition by HI listeners. It might become an even more important factor for challenging listening conditions such as perception of speech degraded by vocoder processing and presented in background noise. In experiment 2, audibility was reduced by testing NH subjects at quiet levels. When the levels were lowered by 15–20 dB compared to experiment 1, the percent correct scores by NH subjects dropped and the reduced performance was more comparable to the performance by HI subjects, for both processed and unprocessed conditions. Hence, audibility had a significant effect on speech recognition by NH subjects as well. However, even at these reduced performance levels, the trends in the performance by the two subject groups still differed. Performance by NH subjects increased up to 12 channels with vowels and up to 10–16 channels with consonants while performance by HI subjects saturated at 8 channels. This observation shows that the inability of the HI subjects to utilize more than eight spectral channels was probably not due to insufficient audibility.

Testing NH subjects at a relatively low presentation level, as it was done in experiment 1, is a reasonable attempt to equalize the listening conditions for both subject groups in loudness, audibility, and comfort. However, by doing so, some potentially detrimental factors due to high presentation levels, such as broadening of AFs (Glasberg and Moore, 2000) or reduced speech recognition in noise (Studebaker *et al.*, 1999), might be overlooked.

To explore the effects of high presentation levels, NH subjects were tested at an absolute presentation level comparable to the presentation level used with the HI group (experiment 3). NH performance did not change significantly compared to experiment 1, for both processed and unprocessed conditions. The performance by NH subjects at the high presentation level was still significantly different when compared to HI listeners; the performance increased up to 12 channels with vowels and 8–12 channels with consonants. Therefore, it is unlikely that the effects observed in the study were due to distortions caused by high presentation levels. However, it is an interesting finding that no effect of AF widening, presumably due to high presentation levels, was observed on performance by NH listeners.

Note that in the present study the audiometric thresholds and the performance intensity functions were measured in quiet, and the presentation levels were adjusted based on these measurements. It is possible that the detection thresholds by HI listeners are considerably higher in noise, yet the present study did not control for this factor.

B. Spectral resolution

The noiseband vocoder was used to systematically degrade the speech stimulus. When background noise was added to the processing, it was observed that HI listeners were not able to use as many vocoder channels as NH listeners for the perception of phonemes. It was hypothesized that this effect was mostly due to reduced frequency resolution of hearing impairment. The inclusion criteria for HI subjects had been determined such that the probability of reduced spectral resolution was maximized; the subjects were selected from a group of patients with flat SNHL ranging from 50 to 60 dB HL, with no OAE and no apparent dead regions.

In most conditions with background noise, the number of channels where the performance by NH listeners saturated was 1.5–2 times larger than the number of channels where the performance by HI listeners saturated. With the assumption that the difference was mostly due to reduced spectral resolution of hearing impairment, this finding implies that the effective average spectral resolution by HI subjects was 1.5–2 times poorer compared to NH subjects in the noisy listening conditions. Even though it is not possible to deduce a conclusion about individual AF shapes for the HI subjects from this observation, it should be noted that this finding is within the ranges of AF broadening observed in SNHL in previous studies. Glasberg and Moore (1986) reported that AFs of impaired ears were broader compared to AFs in normal ears. When expressed in equivalent rectangular bandwidth (ERB), despite the large variation in the data, there were many subjects who showed a broadening in AF bandwidth by a factor of 2. Similarly, Stelmachowicz *et al.* (1985) observed a reduction in Q10 values or the slopes of the low-frequency (LF) end of the psychophysical tuning curves (PTC), another measure for frequency resolution, by a factor of two for many HI subjects compared to NH subjects. These HI subjects had around 50 dB HL or more at the audiometric frequency of 2000 Hz, which is comparable to the hearing loss of the HI subjects in the present study.

Background noise was a major factor for observing the differences in the performance by NH and HI listeners. (1) The difference in the number of channels at the saturation point was observed only in noise. (2) There was a sharper drop in the performance by HI listeners in the unprocessed conditions as the noise level increased. (3) The additional drop in the performance from unprocessed to 40-channel processing condition was generally larger with HI listeners when there was background noise. These findings are consistent with earlier observations that speech perception by HI listeners in quiet is mainly determined by audibility (Festen and Plomp, 1983) while speech perception in noise is more sensitive to suprathreshold deficits (Noordhoek *et al.*, 2000), such as the reduced spectral resolution (Horst, 1987; van Schijndel *et al.*, 2001). In the present study, the difference between NH and HI listeners might have been small in quiet conditions because the audibility was maximized for HI listeners, and the effects of reduced spectral resolution might have showed in noise. One possible explanation for the reduced performance by HI listeners in noise was made by Leek and Summers (1996) who observed that HI listeners with broader AFs needed higher spectral contrast for recognition of speech in noise compared to NH listeners; they suggested that the wide AFs resulted in an internal representation with poorer SNR. The findings of the present study are consistent with this hypothesis; vowel recognition scores by HI listeners at SNR=10 dB, for example, were similar to scores by NH listeners at a higher noise level of SNR =0 dB.

In Experiment 4, the slopes of the vocoder bandpass filters were made shallower in an attempt to simulate broad AFs with NH listeners. Two simulation methods were used: (1) only the carrier filters were made wider, or (2) both the carrier and analysis filters were made wider. The first simulation method did not change the performance significantly. With the second method, the overall performance by NH subjects dropped and the scores were more comparable to the scores by HI listeners. However, the trends by the two subject groups were still different; the scores by NH listeners continued to increase to higher number of channels than HI listeners. Hence, the second simulation method also failed to reproduce the trend of the HI performance.

Fu and Nogaki (2005) used a similar spectral smearing method to simulate the reduced frequency selectivity of CI users with NH subjects. They were able to reproduce the performance by CI users in gated background noise, measured in Speech Reception Threshold (SRT), with NH listeners. However, they measured the SRT only with 4, 8, and 16 channel vocoders, as the main parameter of interest was the frequency of the gated noise. In the present study, the main interest was the change in overall performance as a function of the number of spectral channels, which varied over a wider range of values (2–40) in steady background noise. Similar to Fu and Nogaki (2005), the performance by NH subjects dropped significantly with spectral smearing at 4, 8, and 16 channel conditions. However, when the performance by NH and HI listeners was compared over the entire range of the number of channels used, the trends were significantly different.

The failure to replicate the early saturating characteristic of the HI performance might indicate that the spectral sharpening mechanism of the healthy cochlea is highly effective, to the degree that NH listeners can make use of even minimal spectral contrast of heavily smeared speech stimulus. Baer and Moore (1993), for example, had to simulate AFs six times wider (in ERB) than normal to observe a degradation in speech recognition by NH listeners. A second possibility is that the method used in the present study to simulate braod AFs, where the same cutoff frequencies were used for the filters and only the filter slopes were made shallower, was not entirely adequate. As alternative methods, filter bandwidth could be increased to simulate broad critical bandwidth, as it was suggested by Dreschler and Plomp (1980), or asymmetrical filter shapes could be implemented, to simulate the irregular AF shapes observed with some HI listeners (Sommers and Humes, 1993).

C. Age effects

Some results of the present study were not entirely consistent with the assumption that the effects observed with the vocoder processing and in background noise were due to reduced spectral resolution of hearing impairment. For example, making the vocoder filter skirts shallower failed to replicate the performance by HI listeners with NH listeners, as it was discussed in the previous section. Also, when NH listeners were tested at a high level of 85 dBA, where AFs would presumably be wider (Glasberg and Moore, 2000), the saturation point of the NH performance did not move to a lower number of channels. Moore and Glasberg (2000) observed that the AF shape became asymmetrical with an elevation on the low-frequency side only. It is possible that the AF widening in NH listeners at high levels is to a smaller degree than the AF widening due to cochlear damage (Carney and Nelson, 1983). However, it is also possible that there were additional factors that affected the performance by HI listeners.

Previous research has shown that frequency resolution is generally related to the degree of hearing loss and is independent of age (Peters and Moore, 1992). A correlation with age is usually seen with temporal resolution; older listeners, for example, perform poorly in gap detection task (Snell, 1997; Strouse *et al.*, 1998). This deficit due to aging is thought to be a possible factor for the difficulties that elderly people have in understanding speech in noise (Dubno *et al.*, 1984; Snell and Frisina, 2000). In the present study, HI listeners were older than the NH listeners on average. This might have resulted in potentially different temporal processing abilities between the two subject groups and might have additionally affected the results.

Souza and Boike (2006) repeated the study by Turner *et al.* (1999) with HI subjects of varying age. The performance by older subjects was generally lower compared to younger subjects for all vocoder processing conditions. However, the trend in data was similar; recognition of consonants increased up to 8 channels by all subjects. When the consonant features were analyzed for voicing, manner and place, the only feature that was correlated with age was voicing; older subjects had similar perception of manner and place as younger subjects, but the perception of voicing was poorer by the older subjects.

In the present study, the place cue was expected to be affected most by the spectral degradations. However, in addition to place cue, the transmission of voicing was also affected. The similarity of this finding to observations by Souza and Boike (2006) implies that the HI subjects of the present study might have had deficits in both temporal and spectral processing mechanisms, and the effects of these deficits might have been observed in the perception of different features.

An interesting extension to the present study would be to acquire individual measures from subjects for spectral and temporal resolution, and to correlate these measures with performance by a specific subject. Such a study would provide a more definitive answer to the question if the effects observed in the present study were mainly due to reduced spectral resolution as it was hypothesized.

D. Comparison with implant users

Friesen *et al.* (2001) observed that performance by CI users saturated at a smaller number of channels compared to NH listeners. To explore how the results with implant users compare to results with HI listeners, vowel recognition scores were reproduced from Friesen *et al.* (2001) and superimposed with scores from the present study, for corresponding noise conditions. In Fig. 9, the open and filled circles show the vowel recognition scores from experiment 1 for NH and HI listeners, respectively. The hatched area shows the range of scores, with upper and lower borders defined by best and worst performance by CI users, respectively. The solid line shows the average performance by Nucleus-22 implant users. The scores were measured using the same vowel stimuli in both studies, and the scores with implant users were also corrected for chance level for consistency.

The combined results in Fig. 9 show that at SNR =0 dB, the performance by CI and HI listeners similarly saturated at smaller number of channels (i.e., 7 and 8, respectively) compared to NH listeners (12 channels). The average CI performance was considerably lower than average NH and HI performance, but the best CI performance was similar to the average HI performance. These observations are consistent with results reported by Henry *et al.* (2005) that showed that the frequency resolution with HI listeners was poorer than NH listeners and better than CI listeners, and only the best implant users performed at levels similar to HI listeners.

The poor peformance by implant users, compared to NH and HI listeners, show that there are probably many factors that affect speech recognition by CI users. After all, electric

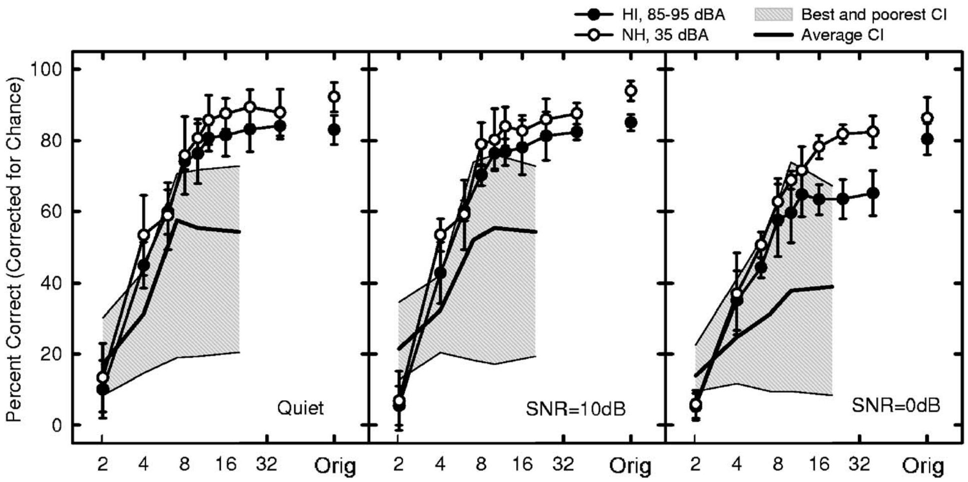

FIG. 9. Average vowel recognition scores by NH and HI listeners presented with average scores by CI users replicated from the study by Friesen *et al.* (2001). From left to right panels the background noise increases. The open and filled symbols show the performance by NH and HI listeners, respectively. The solid line shows the average scores by Nucleus-22 implant users. The hatched area shows the range of the scores by CI subjects.

hearing works on an entirely different mechanism than the acoustic hearing. On the other hand, there was at least one setting (SNR=0 dB) of the present study that produced similar trend in performance by HI and CI listeners. This similarity can be interpreted that an auditory system with inherently reduced frequency resolution, either due to a loss in the peripheral nonlinear mechanism or maybe even due to deficits in the central auditory system, might have a limiting effect on performance similar to CI users whose performance is believed to be limited mainly due to channel interactions. On a more speculative note, one can hypothesize that there might be such factors in addition to channel interactions that are further degrading the CI performance. However, the results of the current study did not completely support this hypothesis as the effects with HI listeners were seen only in noise and there were indications that the performance by HI listeners might have been additionally affected by factors such as reduced temporal resolution, which is usually not observed with implant users.

ACKNOWLEDGMENTS

The author would like to thank Qian-Jie Fu for experimental software CAST, John Galvin for help with the experimental setup and general discussions, House Ear Clinic for help with subject recruitment, the CARE Center for lending the OAE equipment, Sandy Oba for help with screening patients, and José Alcántara, Monita Chatterjee, and Anastasios Sarampalis for valuable suggestions. Funding was provided by a NOHR grant.

American National Standards Institute (**1997**). Methods for Calculation of the Speech Intelligibility Index, S35.

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

Baker, R. J., and Rosen, S. (**2002**). "Auditory filter nonlinearity in mild/moderate hearing impairment," J. Acoust. Soc. Am. **111**, 1330–1339.

Bierer, J. A., and Middlebrooks, J. C. (**2002**). "Auditory cortical images of cochlear-implant stimuli: Dependence on electrode configuration," J. Neurophysiol. **87**, 478–492.

Bierer, J. A., and Middlebrooks, J. C. (**2004**). "Cortical responses to cochlear implant stimulation: Channel interactions," J. Assoc. Res. Otolaryngol. **5**, 32–48.

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

Carney, A. E., and Nelson, D. A. (**1983**). "An analysis of psychophysical tuning curves in normal and pathological ears," J. Acoust. Soc. Am. **73**, 268–278.

Cazals, Y., Pelizzone, M., Kasper, A., and Montandon, P. (**1990**). "Multichannel cochlear implant patients with different open speech understanding show some similar basic psychophysical results," Acta Oto-Laryngol. **S469**, 150–155.

Chatterjee, M., and Shannon, R. V. (**1998**). "Forward masked excitation patterns in multielectrode cochlear implants." J. Acoust. Soc. Am. **103**, 2565–2572.

Dreschler, W. A., and Plomp, R. (**1980**). "Relation between psychopchysical data and speech perception for hearing-impaired subjects," J. Acoust. Soc. Am. **68**, 1608–1615.

Dubno, J. R., Dirks, D. D., and Morgan, D. E. (**1984**). "Effects of age and mild hearing loss on speech recognition in noise," J. Acoust. Soc. Am. **76**, 87–96.

Festen, J. M., and Plomp, R. (**1983**). "Relations between auditory functions in impaired hearing," J. Acoust. Soc. Am. **73**, 652–662.

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.

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

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

Garnham, C., O'Driscoll, M., Ramsden, R., and Saeed, S. (**2002**). "Speech understanding in noise with a Med-El 40+ cochlear implant using reduced channel sets," Ear Hear. **23**, 540–552.

Glasberg, B. R., and Moore, B. C. J. (**1986**). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," J. Acoust. Soc. Am. **79**, 1020–1033.

Glasberg, B. R., and Moore, B. C. J. (**2000**). "Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise," J. Acoust. Soc. Am. **108**, 2318–2328.

Gstoettner, W. K., Adunka, O., Franz, P., Hamzavi, J., Plenk, H., Susani, M., Baumgartner, W., and Kiefer, J. (**2001**). "Perimodiolar electrodes in cochlear implant surgery," Acta Oto-Laryngol. **121**, 216–219.

Henry, B., Turner, C. W., and Behrens, A. (**2005**). "Spectral peak resolution and speech recognition in quiet: Normal-hearing, hearing-impaired, and cochlear implant listeners," J. Acoust. Soc. Am. **118**, 1111–1121.

Hillenbrand, J., Getty, L., Clark, M., and Wheeler, K. (**1995**). "Acoustic characteristics of American English vowels," J. Acoust. Soc. Am. **97**, 3099–3111.

Hornsby, B. W. Y., Trine, T. D., and Ohde, R. N. (**2005**). "The effects of high presentation levels on consonant feature transmission," J. Acoust. Soc. Am. **118**, 1719–1729.

Horst, J. W. (**1987**). "Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects," J. Acoust. Soc. Am. **82**, 874–885.

Kwon, B. J., and van den Honert, C. (**2006**). "Effect of electrode configuration on psychophysical forward masking in cochlear implant listeners," J. Acoust. Soc. Am. **119**, 2994–3002.

Leek, M., and Summers, V. (**1996**). "Reduced frequency selectivity and the preservation of spectral contrast in noise," J. Acoust. Soc. Am. **100**, 1796–1806.

Miller, G. A., and Nicely, P. E. (**1955**). "An analysis of perceptual confusions among some English consonants," J. Acoust. Soc. Am. **27**, 338–352.

Moore, B. C. J. (**1996**). "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," Ear Hear. **17**, 133–160.

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcántara, J. I. (**2000**). "A test for diagnosis of dead regions in the cochlea," Br. J. Audiol. **34**, 205–224.

Noordhoek, I. M., Houtgast, T., and Festen, J. M. (**2000**). "Measuring the threshold for speech reception by adaptive variation of the signal bandwith. II. Hearing-impaired listeners," J. Acoust. Soc. Am. **107**, 1685–1696.

Norton, S. J. (**1992**). "Cochlear function and otoacoustic emissions," Semin. Hear. **13**, 1–14.

Oxenham, A. J., and Plack, C. J. (**1997**). "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **101**, 3666–3675.

Peters, R. W., and Moore, B. C. J. (**1992**). "Auditory filter shapes at low center frequencies in young and elderly hearing-impaired subjects," J. Acoust. Soc. Am. **91**, 256–266.

Pfingst, B. E., Zwolan, T. A., and Holloway, L. A. (**1997**). "Effects of stimulus configuration on psychophysical operating levels and on speech recognition with cochlear implants," Hear. Res. **112**, 247–260.

Pfingst, B. E., Franck, K. H., Xu, L., Bauer, E. M., and Zwolan, T. A. (**2001**). "Effects of electrode configuration and place of stimulation on speech perception with cochlear prostheses," J. Assoc. Res. Otolaryngol. **2**, 87–103.

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

Shannon, R. V. (**1983**). "Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction," Hear. Res. **12**, 1–16.

Shannon, R. V., Zeng, F-G., Kamath, V., Wygonski, J., and Ekelid, M. (**1995**). "Speech recognition with primarily temporal cues," Science **270**, 303–304.

Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E., and Wang, X. (**1999**). "Consonant recordings for speech testing," J. Acoust. Soc. Am. **106**, L71–L74.

Snell, K. B. (**1997**). "Age-related changes in temporal gap detection," J. Acoust. Soc. Am. **101**, 2214–2220.

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.

Sommers, M. S., and Humes, L. E. (**1993**). "Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listeners," J. Acoust. Soc. Am. **93**, 2903–2914.

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.

Stelmachowicz, P. G., Jesteadt, W., Gorga, M. P., and Mott, J. (**1985**). "Speech perception ability and psychophysical tuning curves in hearing impaired listeners," J. Acoust. Soc. Am. **77**, 620–627.

Stickney, G. S., Loizou, P. C., Mishra, L. N., Assman, P. F., Shannon, R. V., and Opie, J. M. (**2006**). "Effects of electrode design and configuration on channel interactions," Hear. Res. **211**, 33–45.

Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (**1998**). "Temporal processing in the aging auditory system," J. Acoust. Soc. Am. **104**, 2385–2399.

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

ter Keurs, M., Festen, J. M., and Plomp, R. (**1992**). "Effect of spectral envelope smearing on speech reception. I," J. Acoust. Soc. Am. **91**, 2872–2880.

ter Keurs, M., Festen, J. M., and Plomp, R. (**1993**). "Effect of spectral envelope smearing on speech reception. II," J. Acoust. Soc. Am. **93**, 1547–1552.

Throckmorton, C. S., and Collins, L. M. (**1999**). "Investigation of the effects of temporal and spatial interactions on speech-recognition skills in cochlear implant subjects," J. Acoust. Soc. Am. **105**, 861–873.

Turner, C. W., and Henn, C. C. (**1989**). "The relation between vowel recognition and measures of frequency resolution," J. Speech Hear. Res. **32**, 49–58.

Turner, C. W., Souza, P. E., and Forget, L. N. (**1995**). "Use of temporal envelope cues in speech recognition by normal and hearing-impaired listeners," J. Acoust. Soc. Am. **97**, 2568–2576.

Turner, C. W., Chi, S.-L., and Flock, S. (**1999**). "Limiting spectral resolution in speech for listeners with sensorineural hearing loss," J. Speech Lang. Hear. Res. **42**, 773–784.

van Schijndel, N. H., Houtgast, T., and Festen, J. M. (**2001**). "Effects of degradation of intensity, time, or frequency content on speech intelligibility for normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **110**, 529–542.

van Tasell, D. J., Greenfield, D. G., Logemann, J. J., and Nelson, D. A. (**1992**). "Temporal cues for consonant recognition: Training, talker generalization, and use in evaluation of cochlear implants," J. Acoust. Soc. Am. **92**, 1247–1257.

van Tasell, D. J. (**1993**). "Hearing loss, speech, and hearing aids," J. Speech Hear. Res. **36**, 228–244.

Wilson, B., Finley, C., Laeson, D., Wolford, R., Eddington, D., and Rabinowitz, W. (**1991**). "Better speech recognition with cochlear implants," Nature (London) **352**, 236–238.

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.