Open access • Journal Article • DOI:10.1121/1.1944567

Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. — Source link

Belinda A. Henry, Christopher W. Turner, Amy Behrens

Institutions: University of Wisconsin-Madison

Published on: 04 Aug 2005 - Journal of the Acoustical Society of America (Acoustical Society of America)

Topics: Noise

Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners

Belinda A. Henry[a)]
Department of Communicative Disorders, University of Wisconsin—Madison, Madison, Wisconsin 53706

Christopher W. Turner
Department of Speech Pathology and Audiology, The University of Iowa, Iowa City, Iowa 52242

Amy Behrens
Department of Speech Pathology and Audiology, The University of Iowa, Iowa City, Iowa 52242

(Received 10 May 2004; revised 8 May 2005; accepted 9 May 2005)

Spectral peak resolution was investigated in normal hearing (NH), hearing impaired (HI), and cochlear implant (CI) listeners. The task involved discriminating between two rippled noise stimuli in which the frequency positions of the log-spaced peaks and valleys were interchanged. The ripple spacing was varied adaptively from 0.13 to 11.31 ripples/octave, and the minimum ripple spacing at which a reversal in peak and trough positions could be detected was determined as the spectral peak resolution threshold for each listener. Spectral peak resolution was best, on average, in NH listeners, poorest in CI listeners, and intermediate for HI listeners. There was a significant relationship between spectral peak resolution and both vowel and consonant recognition in quiet across the three listener groups. The results indicate that the degree of spectral peak resolution required for accurate vowel and consonant recognition in quiet backgrounds is around 4 ripples/octave, and that spectral peak resolution poorer than around 1–2 ripples/octave may result in highly degraded speech recognition. These results suggest that efforts to improve spectral peak resolution for HI and CI users may lead to improved speech recognition. © *2005 Acoustical Society of America.* [DOI: 10.1121/1.1944567]

PACS number(s): 43.71.Ky, 43.66.Ts, 43.66.Sr, 43.71.Es [KWG] Pages: 1111–1121

I. INTRODUCTION

Speech recognition in listeners with sensorineural hearing loss varies widely, both among those with impaired acoustic hearing [hearing impaired (HI) listeners] and among those using cochlear implants (CI listeners). While some HI and CI listeners achieve high levels of audition-alone open-set speech recognition, at the other end of the range are some listeners who must rely on supplementary visual cues in order to understand speech. The loss of absolute sensitivity in HI listeners is the primary factor affecting speech perception, and amplification via the use of hearing aids compensates for this to some extent. However, for those HI listeners with hearing losses in the moderate to profound range, audibility does not account for the entire deficit in speech perception (e.g., Ching *et al.*, 1998; Dubno *et al.*, 1989; Hogan and Turner, 1998; Humes *et al.*, 1986; Pavlovic, 1984; Pavlovic *et al.*, 1986; Skinner, 1980). In these cases, variability in performance is likely to be related not only to the audibility of speech cues, but also to abnormalities in the perceptual analysis of sound at suprathreshold levels. In CI listeners, it is this latter factor that is thought to be related to performance variability. Acoustic signals are transformed by the speech processor, and speech cues are represented in the pattern of electrical stimulation across the electrode array. Audibility is not the primary factor contributing to performance variability, since the audibility of conversational-level speech is determined primarily by the input dynamic range and sensitivity setting of the speech processor (provided that the electrical stimulation levels and sensitivity control are set optimally). Rather, it is the ability to extract speech cues from the audible patterns of electrical stimulation that is the most important factor limiting performance in CI listeners.

One perceptual factor that is likely to limit speech perception in both HI and CI listeners is reduced spectral resolution. Accurate speech recognition depends partly on the ability to perceive the spectral shapes of speech sounds, and, in particular, to identify the frequencies of spectral peaks. In normal hearing (NH) listeners, the frequency selectivity of the auditory system is thought to underlie the process of resolving spectral peaks in the speech signal. Impaired frequency selectivity has been demonstrated in HI listeners in both physiological (e.g., Dallos *et al.*, 1977; Liberman and Dodds, 1984) and psychophysical studies (e.g., Dubno and Dirks, 1989; Glasberg and Moore, 1986; Trees and Turner, 1986; Wightman *et al.*, 1977). The bandwidth of auditory filters has been shown to be up to three to four times greater than normal in listeners with cochlear hearing loss (Glasberg and Moore, 1986). Spectral resolution is also reduced in CI listeners, although for different underlying reasons than in HI listeners. Multichannel CIs replace the peripheral frequency selectivity of the normal auditory system with multiple intra-

[a)]Electronic mail: bahenry@wisc.edu

cochlear electrodes. Spectral cues are coded by resolving the frequency components of the signal using bandpass filtering, and mapping the outputs of these bands onto the intracochlear electrodes in a tonotopically assigned manner. Spectral resolution is limited in CI listeners by the number of stimulating channels (currently between 6 and 22, depending on the device and speech processing strategy) and the ability of the individual to resolve the spectral cues that are provided.

What is the effect of impaired spectral resolution on speech recognition? One might hypothesize that if spectral resolution is reduced, either due to an impaired auditory system or the use of a CI, the spectral envelope may be "blurred," making it difficult for a listener to identify the frequency locations of spectral peaks in speech. It is of theoretical importance, as well as significant practical importance in designing improved hearing aids and CIs and optimizing these devices for individuals, to determine the relationship between spectral peak resolution and speech recognition, and the degree of spectral peak resolution that is required for accurate speech recognition. In other words, a question of particular importance is as follows: What is the minimum requirement for spectral peak resolution, below which speech recognition becomes degraded? These questions have been investigated by numerous authors using at least two different approaches.

The first approach involves manipulating the spectral resolution available in the speech signal and examining the effects of this processing on speech recognition both in NH listeners as well as in HI and CI listeners. The results of studies on the effects of simulated reduced spectral resolution in NH listeners indicate that speech recognition in quiet is highly resistant to degraded spectral resolution. Several authors have investigated the effects of simulated broadened auditory filters by measuring speech recognition in NH listeners under conditions of spectral smearing. Baer and Moore (1993) and ter Keurs *et al.* (1992) showed that simulating auditory filters up to six times broader than those of NH listeners has little effect on speech recognition in quiet. Furthermore, Boothroyd *et al.* (1996) showed that in order to reduce phoneme recognition in quiet by 50%, the spectral information needed to be smeared by as much as 1400 Hz. Simulations of CI processing, in which multiple bands of speech-modulated noise are presented to NH listeners, show high levels of speech recognition for NH listeners in quiet with between 4 and 12 spectral channels, depending on the degree of difficulty of the speech materials (Dorman *et al.*, 1997; Friesen *et al.*, 2001; Shannon *et al.*, 1995; Turner *et al.*, 1995). These findings further indicate that fine spectral resolution is not required for speech recognition in quiet.

While the CI simulation studies indicate that CIs can ideally present sufficient spectral detail for accurate speech recognition in quiet, studies on the effect of the number of channels on speech recognition in CI recipients indicate that the effective number of channels perceived by these listeners is lower than the physical number of channels provided. CI users show an asymptote in speech recognition on average across listeners with between two and seven channels, depending on the degree of difficulty of the speech material presented (Dorman and Loizou, 1997; Fishman *et al.*, 1997; Friesen *et al.*, 2001). Furthermore, the effect of the number of channels varies widely, with better-performing CI listeners generally able to use more channels than those with poorer overall performance (Friesen *et al.*, 2001). The effects of limiting the spectral resolution provided in the speech signal on performance in HI listeners has also been investigated (Turner *et al.*, 1999). The HI listeners in that study performed equivalently to NH listeners for single band speech. However, as the number of bands of speech-modulated noise presented was increased, the performance of HI listeners did not increase to the same extent as for NH listeners, indicating that some HI listeners cannot utilize all the spectral information in speech.

The second approach is to attempt to relate performance on psychophysical measures of spectral resolution to speech recognition in HI and CI listeners. If reduced frequency resolution is associated with poorer speech recognition, a statistical relationship between these two measures may be expected. However, in HI listeners strong correlations between speech recognition in quiet and frequency selectivity, as measured in psychoacoustic masking experiments for instance, have been difficult to establish and findings vary across studies (Dreschler and Plomp, 1980, 1985; Festen and Plomp, 1983; Lutman and Clark, 1986; Glasberg and Moore, 1989; Stelmachowicz *et al.*, 1985; Tyler *et al.*, 1982). In studies where correlations were found, it was difficult to separate the roles of frequency selectivity and audibility in speech recognition since both frequency selectivity and speech recognition were correlated with absolute hearing thresholds. Correlations between speech recognition and frequency selectivity were often reduced or eliminated after the effect of absolute threshold was statistically partialled out. In CI listeners, place of stimulation perception, as determined using psychophysical measures such as electrode discrimination and pitch ranking, is generally assumed to underlie to some extent the ability to resolve the spectral aspects of the speech signal. While several authors have reported a relationship between speech recognition and place of stimulation perception (Collins *et al.*, 1997; Donaldson and Nelson, 2000; Nelson *et al.*, 1995; Throckmorton and Collins, 1999), Zwolan *et al.* (1997) showed no correlation between these two measures, and Henry *et al.* (2000) showed a correlation for the low- to mid-frequency regions only, and only when there was random level variation between stimuli. These results generally indicate that those CI listeners who are more sensitive to the place of stimulation in the cochlea, particularly in the presence of random level variation, are better at recognizing speech.

The traditional measures of frequency resolution in HI listeners and place of stimulation perception in CI listeners, which typically require a listener to detect a signal in the presence of a masker (HI listeners) or to discriminate between stimulation on different electrodes activated individually (CI listeners), are indirect measures of spectral peak resolution for complex broadband acoustic signals. The distinct methodologies that have been employed in these studies do not allow the comparison of spectral peak resolution abilities across NH, HI, and CI listeners. Consequently, the gen-

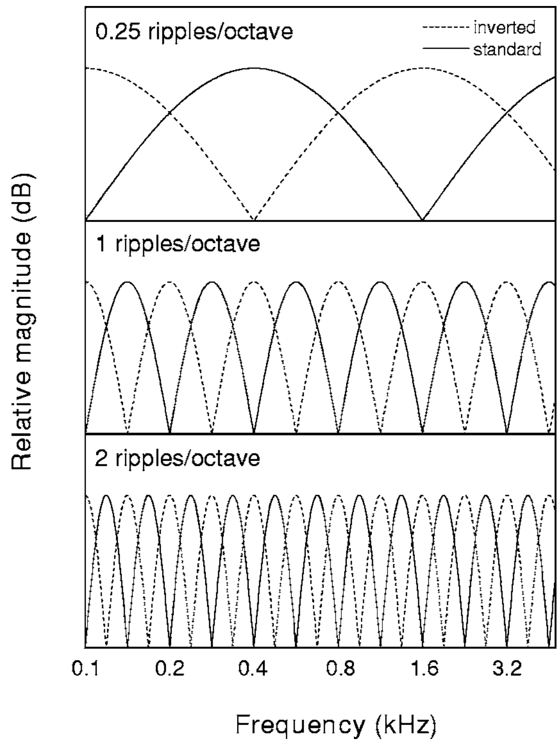

FIG. 1. Rippled noise spectra. Standard and inverted peak positions for ripple frequencies of 0.25, 1 and 2 ripples/octave are shown.

eral relationship between spectral peak resolution and speech recognition and the degree of spectral peak resolution that is required for accurate speech recognition currently remains somewhat unclear.

A more direct measure of spectral peak resolution for acoustic signals was recently developed by Henry and Turner (2003) and applied to CI listeners. The method is based on the "ripple phase reversal test" (Supin *et al.*, 1994). In Henry and Turner (2003) the stimuli were rippled noise signals, which are broadband noise signals with spectral ripples spaced on a linear scale. The task involves discriminating between two rippled noise stimuli in which the frequency positions of the peaks and valleys are interchanged. The ripple spacing is varied (with the ripple depth held constant), and the minimum ripple spacing at which a reversal in peak and trough positions can be detected is determined as the threshold for spectral peak resolution. Examples of rippled noise stimuli (with a logarithmic spacing of ripples in this case; see below) are shown in Fig. 1. This test is hypothesized to provide a direct measure of the ability of listeners to perceive the frequency locations of spectral peaks in a broadband acoustic signal. The results showed a significant relationship between spectral peak resolution and vowel recognition in CI listeners, indicating that listeners who can resolve more closely spaced peaks are better at recognizing vowels. This measure has potential applications in predicting and optimizing speech recognition in CI listeners. Furthermore, it is hypothesized that this test may also be applicable to listeners with acoustic hearing, and may therefore provide a measure of spectral peak resolution in HI listeners. As such, this test provides the opportunity to directly compare spectral peak resolution abilities among listeners with acoustic hearing (normal or impaired) and listeners with electric hearing.

In the present study, the spectral peak resolution test was adapted for the investigation of spectral peak resolution in NH and HI listeners, and for a further investigation of this ability in CI listeners. The spectral peak resolution test was implemented in this study using logarithmically spaced ripples, instead of using linear-spaced ripples, as in Henry and Turner (2003). There are two reasons for using a logarithmic spacing of ripples. First, it is hypothesized that such a spacing would more closely approximate the properties of the normal auditory system as well as the acoustics of speech, and may therefore be more strongly correlated with speech recognition. Second, since it is thought that the processing of linear rippled noises may involve a time-domain waveform analysis in the acoustic auditory system (e.g., Fay *et al.*, 1983; Yost *et al.*, 1996), log-rippled spectra are better suited to the specific examination of spectral peak resolution in listeners with acoustic hearing.

The modified spectral peak resolution test was used to investigate the differences in spectral peak resolution between and among NH, HI, and CI listeners. In addition, the possible relationship between spectral peak resolution and speech recognition both across the NH, CI, and HI listener groups, and within each of the CI and HI listener groups was examined. The present experiments, by measuring these relations across the wide range of listeners, can therefore test whether the ability to perceive the locations of spectral peaks is a requirement for speech recognition in general, and if there is some minimum requirement in this ability, below which speech perception is highly degraded.

II. METHODS

A. Subjects

Three subject groups participated in this study: (1) NH listeners, (2) HI listeners, and (3) CI listeners. All participants were native American English speaking adults. There were 12 young adult NH subjects. Normal hearing was defined as having pure-tone air conduction thresholds ≤15 dB HL at octave frequencies from 125 to 8000 Hz in the tested ear.

Thirty-two HI listeners ranging in age from 29 to 83 years participated. The hearing losses were diagnosed as sensorineural (and assumed to be of cochlear origin) based on the lack of an air-bone gap and tympanograms consistent with normal middle ear function. The ear with the better pure tone thresholds was selected as the test ear. The degree of hearing loss ranged from mild to profound, and the audiometric configurations (flat or sloping) varied across the HI listeners. Pure-tone thresholds for the test ear for each subject, along with the ages of each subject, are shown in Table I. In Table I, the downward-pointing arrows indicate that the thresholds were higher than could be measured with the audiometer.

The 23 CI subjects were users of the Cochlear Ltd. Nucleus 24M (CI24M) or Nucleus 24 Contour (CI24R) implant and had a minimum of 6 months experience with their

TABLE I. Individual subject details: Hearing impaired subjects. Audiometric thresholds for the test ear are shown in dB HL.

Subject	Age (yrs)	Frequency (Hz) 250	500	1000	1500	2000	3000	4000	6000	8000
HI1	55	35	40	50		50		40		50
HI2	69	25	20	25	35	55		65		85
HI3	55	10	10	25	30	45	60	70		55
HI4	64	10	15	20	55	85		80		75
HI5	75	40	60	50		35	45	65		65
HI6	47	15	15	45		35	20	15		15
HI7	81	20	25	25	25	65	70	70		80
HI8	55	15	10	15	60	70		70	85	95
HI9	29	60	70	100		100		↓		↓
HI10	71	40	40	55		45		40		65
HI11	63	60	55	45		40		45		60
HI12	59	35	35	30	45	55		55	80	85
HI13	76	15	15	25	25	45		60	60	80
HI14	69	65	60	65		70		55		85
HI15	57	35	25	20		25	60	65		55
HI16	61	15	20	35	50	60		70		65
HI17	61	35	30	55		60	90	105		↓
HI18	75	25	35	40		55	75	80		95
HI19	75	40	35	50		60		65		75
HI20	75	45	45	45		45		50		65
HI25	62	10	5	15	30	35	70	80		90
HI26	53	35	50	45		45		55		55
HI27	79	35	35	40		45		55		55
HI28	79	50	60	75		75		70		65
HI29	65	55	50	50		45		55		60
HI30	61	20	30	30	35	50		55		50
HI31	79	35	40	35		70		75		90
HI32	83	35	40	35		40	45	65		70
HI33	79	20	30	35		50	55	55	45	50
HI34	71	50	50	60		80	105	↓		↓
HI35	73	25	65	100		95		85		75
HI36	76	15	25	40		55	65	75		70

implant. Nineteen subjects used the CI24M implant, which has 22 intracochlear and 2 extracochlear electrodes, while the remaining subjects used the CI24R implant, which has a preformed (curved) perimodiolar electrode array instead of a straight array, and is designed to be positioned adjacent to the modiolar wall, decreasing the distance to the target neurons. Three subjects used the Continuous Interleaved Sampling (CIS) strategy (Wilson *et al.*, 1991), 5 used the SPEAK strategy (Seligman and McDermott, 1995), and 15 used the Advanced Combination Encoder (ACE) strategy (Skinner *et al.*, 2002; Vandali *et al.*, 2000). In the CIS strategy implemented with the Nucleus device, the amplitude envelope is estimated within each of typically 6–12 channels during each stimulation period. These amplitudes are converted to electrical stimulation levels, and stimulus pulses representing each band are presented sequentially on the associated electrodes at a rate between 740 and 2400 pulses per second (pps)/channel. SPEAK and ACE are both "peak-picking" strategies that estimate the amplitude envelope in up to 20 (SPEAK) or 22 (ACE) channels, each assigned in a tonotopic order to an equal number of implanted electrodes. In each analysis cycle, the channels with the largest amplitudes (maxima) are selected, and stimulus pulses are then presented sequentially on the associated electrodes. The number of maxima selected is 6 on average for SPEAK and between 1 and 20 for ACE (typically between 8 and 12), and the rate of stimulation is approximately 250 pps/channel for SPEAK and is between 250 and 2400 pps/channel (limited by the maximum rate of 14 400 pps across all channels) for ACE. Individual subject details are shown in Table II, including the parameters used for each subject's map. Also shown in Table II are audition-alone word recognition scores for CNC words, measured in the University of Iowa Cochlear Implant Clinic during the most recent speech processor mapping session with the subject's clinical map.

B. Stimuli

Speech recognition was assessed using vowel and consonant stimuli. The consonant test used a closed-set 16-alternative identification paradigm for consonants presented in an /a/-consonant-/a/ context (Turner *et al.*, 1995). The tokens were produced by four talkers (2 female and 2 male). Each talker produced one token of each of the /aCa/syllables,

TABLE II. Individual subject details: Cochlear implant subjects. Prog.=progressive; ACE used 8 maxima unless otherwise noted.

Subject	Age (yrs)	Duration of profound deafness (yrs)	CI experience (yrs)	Etiology	Implant type, processor type	Processing strategy, # maxima	Pulse rate (pps/ch), # channels	CNC word score (% correct)	Average dynamic range (dB)
CI1	74	12	6	Infection	CI24M	ACE	720, 22	86	8.6
CI2	47	13	3	Congenital, prog.	CI24M	ACE	900, 22	39	12.5
CI3	64	4	5	Unknown	CI24M	ACE	720, 18	56	9.7
CI5	73	8	4	Congenital, prog.	CI24M	ACE	900, 20	64	7.6
CI6	73	1	2	Meniere's disease	CI24M	SPEAK	250, 18	54	4.9
CI7	44	0.5	3	Autoimmune disease	CI24M	ACE	1200, 22	72	15.5
CI10	75	25	5	Congenital, prog.	CI24M	ACE	720, 20	68	7.7
CI11	77	40	3	Unknown	CI24M	CIS	900, 6	4	6.9
CI13	49	2	2	Unknown	CI24M	CIS	2400, 6	74	12.0
CI14	55	5	4	Unknown	CI24M	ACE, 10	1200, 22	18	4.0
CI15	81	3	4	Unknown	CI24M	SPEAK	250, 19	54	4.9
CI16	37	2	3	Unknown	CI24R	ACE, 12	720, 22	42	8.8
CI18	57	36	6	Unknown	CI24M	ACE, 12	1200, 20	50	9.0
CI19	79	0.5	5	Viral infection	CI24M	SPEAK	250, 20	66	2.5
CI20	47	7	4	Unknown	CI24M	ACE	900, 20	22	16.5
CI22	63	0.3	3	Infection	CI24M	ACE	720, 22	82	10.9
CI23	75	8	3	Unknown	CI24M	SPEAK	250, 20	42	4.9
CI24	85	11	2	Unknown	CI24R	SPEAK	250, 18	24	5.8
CI25	76	10	4	Unknown, prog.	CI24M	ACE	900, 22	68	11.6
CI26	62	1	3	Meniere's disease	CI24M	ACE	720, 22	54	9.5
CI27	47	28	0.5	Infection	CI24R	ACE	900, 22	64	19.2
CI28	41	3	2	Hereditary	CI24R	ACE	900, 22	84	9.5
CI29	49	8	3	Unknown, prog.	CI24M	CIS	900, 6	58	7.7

and each token was repeated in random order 3 times in the test, for a total of 192 test items. Vowel recognition was measured using a closed-set 12-alternative identification procedure. Medial vowel tokens produced by 10 male and 10 female talkers were selected from the materials recorded by Hillenbrand *et al.* (1995), and presented in a /h/-vowel-/d/ context, for a total of 240 test items. The speech stimuli were stored in digital form on a Macintosh G4 computer.

Rippled noise stimuli of 100–5000 Hz bandwidth and with peak-to-valley ratios of approximately 30 dB were synthesized on an Apple Macintosh G4 computer by algebraically summing 200 pure-tone frequency components with amplitudes determined by a sinusoidal envelope with ripples spaced on a logarithmic frequency scale. The starting phases of the individual frequency components were randomized for each stimulus to avoid fine structure pitch cues that may be perceptible to listeners. The frequency of the spectral envelope of the stimulus complex was varied in 14 steps: 0.125, 0.176, 0.250, 0.354, 0.500, 0.707, 1.000, 1.414, 2.000, 2.828, 4.000, 5.657, 8.000, and 11.314 ripples per octave (ripples/octave). The spectral envelope phase of the stimulus complex was set to zero at the low-frequency edge of the complex for the standard (reference) stimulus, and the inverted (test) stimulus had a reversed phase. Examples of standard and inverted 0.25, 1 and 2 ripples/octave rippled noise spectra are shown in Fig. 1. The stimuli were of 500 ms duration and had 150 ms rise/fall times, and were shaped with a filter that approximated the long-term speech spectrum (Byrne *et al.*, 1994). The overall levels of the rippled noise sound files were then approximately equalized.

C. Procedures

All subjects were tested in a double-walled sound treated room. The speech and rippled noise stimuli were output via custom software routines through a 16-bit DigiDesign (Audio Media III) digital-to-analog converter at a sampling rate of 44.1 kHz. These stimuli were presented to NH and HI subjects monaurally through Sennheiser HD 25-SP1 circumaural headphones. The presentation level for both speech and rippled noise stimuli was 65 dB SPL for the normal-hearing listeners. Stimuli were presented to HI subjects through an analog high-pass emphasis spectrum shaper (Altec-Lansing 1753), which provided approximately 20 dB of relative gain with a transition slope of 40 dB/octave, starting at 1000 Hz. High-frequency emphasis was not provided to either the NH or CI subjects. The presentation level was set on an individual basis for each of the HI listeners at the highest possible level that was acceptable for each subject, as determined in pilot test sessions. The aim was to optimize the audibility of the signals across the wide range of frequencies for the HI listeners, although it should be noted that the high-pass emphasis would not have been sufficient to provide signal audibility in the high frequencies for some subjects. The chosen level for the speech materials was then also used for the rippled noise stimuli.

The speech and rippled noise stimuli were presented to the CI subjects using the SPrint speech processor in the free field, positioned approximately 1 m from a loudspeaker (Cerwin-Vega Model E712), at an average level of 65 dB SPL. The laboratory SPrint speech processor was programmed with the clinical map used by each subject (see Table II). While it seems highly likely that spectral resolution may be affected by electrode array design, speech processing strategy, and processing parameters, our purpose in this study was not to directly investigate these effects. Instead, these potential sources of variability in spectral resolution across CI listeners could be exploited to test if spectral resolution abilities were predictive of speech recognition. Therefore, various processing strategies and electrode designs were specifically included in this study in order to assess the relationship between spectral peak resolution and speech recognition with the everyday map used by each subject. In all cases, stimulus pulses of 25 μs duration were presented using the monopolar M1+2 electrode configuration, where current flows between the active intracochlear electrode and both extracochlear electrodes. Prior to commencing the experiment, threshold (T) and comfortably loud (C) levels were measured for each electrode, using standard clinical procedures. In order to minimize as much as possible individual variation in the audibility of acoustic signals, the same speech processor sensitivity (set to 8) was used for all subjects. The sensitivity was set so that peaks in the stimulus resulted in electrical stimulation at approximately 90% of the dynamic range for the 65 dB SPL acoustic input signal. This setting was determined by measuring the speech processor output using the SCILAB (Swiss Cochlear Implant Laboratory software) program (Lai *et al.*, 2003), which records the RF transmissions from the speech processor and provides the current levels for all activated electrodes. Reference was also made to the published sound pressure levels that result in electrical stimulation for the range of sensitivity settings for the SPrint processor (Nucleus Technical Reference Manual, Fig. 3.12).

A single run of the speech tests consisted of 240 trials for vowels and 192 items for consonants. The test items (12 words containing the medial vowels for the vowel test; 16 individual consonants for the consonant test) were displayed as buttons on a touchscreen (MicroTouch). On each trial, a stimulus token was chosen randomly, without replacement, and following the presentation of each token the subject responded by pressing one of the buttons on the touch screen. Two runs (a practice and a test run) of each test were administered. Correct-answer feedback was provided during the practice run only.

Prior to speech testing, training was provided for both the vowel and consonant identification tasks. Subjects were instructed to press the button on the touch screen corresponding to the phoneme they wanted to hear, and five examples of that phoneme (i.e., the phoneme spoken by five different talkers) were then presented. Fifty trials of the training task were conducted (or more as desired by the individual subject) in order to allow the subjects to familiarize themselves with the phonemes and their associated touch screen labels.

Ripple resolution thresholds were determined using a three interval forced-choice adaptive procedure, based on the method developed by Henry and Turner (2003). For each set of three intervals, two intervals contained the standard or reference stimulus, and the test interval, chosen at random, contained the inverted stimulus. There was an interstimulus

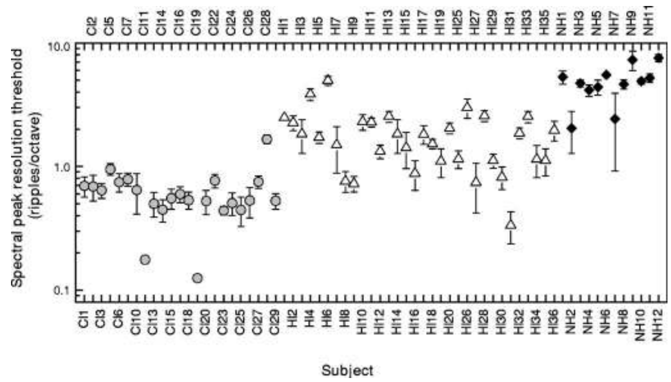

FIG. 2. Thresholds for spectral peak resolution for NH, HI, and CI subjects. Error bars represent ± one standard deviation.

interval of 500 ms. In order to minimize the use of loudness cues, the presentation level was varied randomly from trial to trial within an 8 dB range in 1 dB steps, using a Tucker-Davis Technologies programmable attenuator. Three numerically labeled buttons were displayed on the touch screen, corresponding to the three intervals, and subjects were instructed to press the button corresponding to the interval that sounded "different" (i.e., that contained the test stimulus), ignoring any loudness variation between intervals. Correct answer feedback was provided throughout the experiment. Each test run commenced at a ripple frequency of 0.176 ripples/octave, and the ripple frequency was varied in a two-down, one-up procedure. After each incorrect response the ripple frequency was decreased by a step, and it was increased after two correct responses. This procedure converged on the 70.7% correct point (Levitt, 1971) for ripple resolution. The threshold was estimated for each run as the geometric mean of the ripple frequencies for the final 8 of 12 reversals. Based on previous results using the linear rippled noise stimuli in this laboratory (Henry and Turner, 2003), and pilot testing for these logarithmic rippled noise stimuli, only a few practice runs are necessary to achieve asymptotic performance. Therefore, four practice runs were completed for each subject. Following the practice runs, three test runs were obtained for each subject, and the final threshold value for each subject was recorded as the arithmetic mean of the thresholds across these three test runs.

III. RESULTS

The mean spectral resolution thresholds are shown for each subject in Fig. 2. Higher thresholds (more ripples per octave) indicate a better spectral peak resolution ability. Spectral peak resolution varied from 0.13 to 7.55 ripples/octave across all listeners. NH listeners had the best spectral peak resolution, with an average threshold across listeners of 4.84 ripples/octave and a range of 2.03–7.55 ripples/octave, while CI listeners had the poorest spectral peak resolution, with an average threshold across listeners of 0.62 ripples/octave, and a range of 0.13–1.66 ripples/octave. The average spectral peak resolution threshold of 1.77 ripples/octave for the HI listeners was between those of the NH and the CI

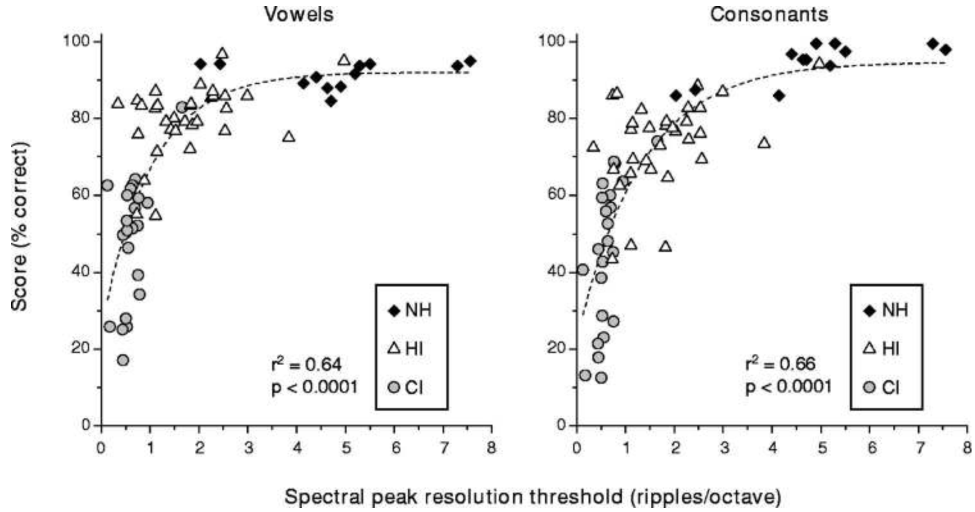

FIG. 3. The relationship between spectral peak resolution and vowel recognition (left panel) and consonant recognition (right panel) across NH, HI, and CI subjects. The dashed curves represent the functions of best fit to the data [Eq. (1)].

groups, and the individual thresholds of 0.33–4.97 ripples/octave essentially spanned the range of thresholds of the NH and CI groups.

The relationship between spectral peak resolution and both vowel and consonant recognition across the three listener types (NH, HI, and CI) is shown in Fig. 3. The following function was found to provide the best fit to both the vowel and consonant data:

$$P = ae^{-S/b} + c, \quad (1)$$

where P is the percent correct score, S is the spectral peak resolution threshold, and a, b, and c are fitting parameters. For the vowel data, $a=-66.88$, $b=1.03$, and $c=92.12$, and for the consonant data, $a=-72.24$, $b=1.33$, and $c=94.76$. Nonlinear regression analysis based on the fitted functions indicated a significant relationship between spectral peak resolution and both vowel recognition ($r^2=0.64, p<0.0001$) and consonant recognition ($r^2=0.66, p<0.0001$). These results suggest that the ability to resolve spectral peaks in a complex acoustic spectrum may be associated with accurate speech recognition.

For vowel recognition, there was an asymptote in performance at a score of approximately 92% correct, which corresponded to a spectral peak resolution threshold of approximately 4 ripples/octave, and for consonant recognition there was an asymptote in performance at a score of approximately 94%, which corresponded to a spectral peak resolution threshold of approximately 4.5 ripples/octave. There was a rapid deterioration in performance for both vowel and consonant recognition when spectral peak resolution fell below approximately 1–2 ripples/octave.

The relationship between spectral peak resolution and speech recognition was also examined for the CI and HI listener groups individually. The relationship between spectral peak resolution and vowel recognition (left panel) and consonant recognition (right panel) for the CI listener group is shown in Fig. 4 and for the HI listener group in Fig. 5. Regression analyses showed a significant moderate linear correlation between spectral peak resolution thresholds and both vowel recognition ($r^2=0.27, p=0.01$) and consonant

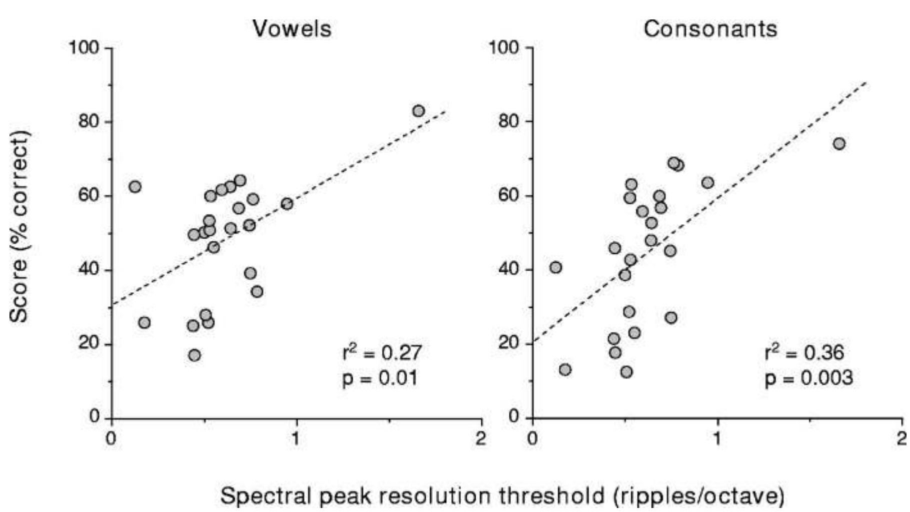

FIG. 4. The relationship between spectral peak resolution and vowel recognition (left panel) and consonant recognition (right panel) for CI subjects. Linear regressions are represented by the dashed lines.

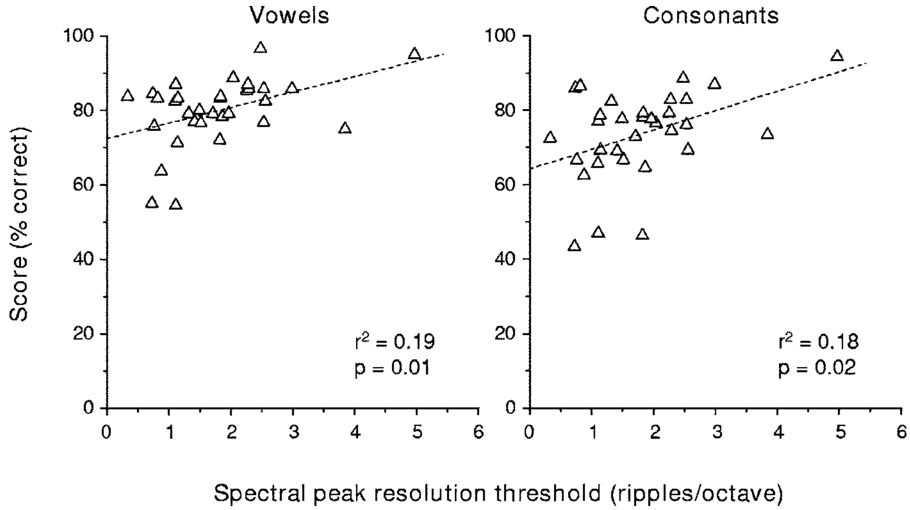

FIG. 5. The relationship between spectral peak resolution and vowel recognition (left panel) and consonant recognition (right panel) for HI subjects. Linear regressions are represented by the dashed lines.

recognition ($r^2=0.36, p=0.003$) for the CI listeners and a significant but quite weak linear correlation between spectral peak resolution thresholds and both vowel recognition ($r^2=0.19, p=0.01$) and consonant recognition ($r^2=0.18, p=0.02$) for the HI listeners.

For the HI listener group, while there was a significant correlation between absolute threshold (calculated as the average of thresholds at 0.5, 1 and 2 kHz) and speech recognition (consonants: $r^2=-0.22$, $p=0.006$; vowels: $r^2=-0.24$, $p=0.004$), there was no correlation between absolute threshold and spectral peak resolution threshold ($r^2=-0.03, p=0.35$). Therefore it seems unlikely that the observed relationship between speech recognition and spectral peak resolution ability was due to underlying relationships between each of these variables and absolute threshold. This finding is in contrast to previous studies that have generally shown a correlation of both frequency selectivity and speech recognition with absolute hearing thresholds in HI listeners, and a reduction or elimination of correlations between frequency selectivity and speech recognition when the effect of absolute threshold was statistically partialed out (see the Introduction). The reason for the lack of a correlation between spectral peak resolution and absolute threshold in this study is not clear. It may suggest that absolute threshold may not be associated with the ability to perform this spectral peak resolution task, possibly due to the fact that the audibility of the rippled noise signals was optimized for individual listeners, or may arise from the fact that broadband stimuli were used to assess spectral peak resolution, where listeners could use the region in which spectral resolution was best, while absolute threshold was averaged across frequency.

For both the CI and HI listeners there were no significant correlations between age and either speech recognition (averaged consonant and vowel score, CI: $r^2=-0.11$, $p=0.12$; HI: $r^2=0.01$, $p=0.55$) or spectral peak resolution threshold (CI: $r^2=-0.17$, $p=0.07$; HI: $r^2=-0.01$, $p=0.56$).

IV. DISCUSSION

The research reported in this study suggests a possible relationship between spectral peak resolution and speech recognition across the NH, HI, and CI listener groups (Fig. 3). The regression analyses between spectral peak resolution and speech recognition across all listeners accounted for 64% of the variance in vowel scores and 66% of the variance in consonant scores, indicating that there may be an underlying dependence for speech recognition in general upon spectral peak resolution. These relationships appear to be stronger than those demonstrated in previous studies that have attempted to link speech recognition and spectral resolution in HI listeners (e.g., Dreschler and Plomp, 1980, 1985; Festen and Plomp, 1983; Lutman and Clark, 1986; Stelmachowicz *et al.*, 1985; Tyler *et al.*, 1982) and speech recognition and place of stimulation perception in CI listeners (e.g., Collins *et al.*, 1997; Donaldson and Nelson, 2000; Henry *et al.*, 2000; Nelson *et al.*, 1995; Throckmorton and Collins, 1999; Zwolan *et al.*, 1997). Stronger relationships in the present study may result from assessing spectral resolution ability in a wide range of individuals across the clinical populations. The spectral peak resolution measure used in this study provides the opportunity to investigate spectral peak resolution in NH, HI, and CI listeners, which has not been possible in the previous studies since the methodologies used to assess spectral resolution in the acoustic hearing (such as psychoacoustic masking measures) and electric hearing (such as electrode discrimination measures) listener groups individually cannot be applied to the assessment of spectral resolution across all groups. While the overall relationship between spectral peak resolution and speech recognition is nonlinear [Fig. 3 and Eq. (1)], reference to the data subsets for the isolated CI and HI groups (Figs. 4 and 5, respectively) shows that the overall nonlinear relationship seen across the listener groups is reduced and is linear. This indicates that restricting the examination of possible relationships between spectral resolution and speech recognition to individual listener groups, as has been done in previous studies, may obscure the overall nonlinear form of the relation and reduce its strength.

What degree of spectral peak resolution is required for accurate speech recognition, and below what degree of spectral peak resolution does speech recognition become highly degraded? While the ability to resolve a higher number of ripples per octave was generally associated with better speech recognition, there was a plateau in multitalker vowel and consonant recognition at around four ripples/octave (Fig. 3). This indicates that spectral peak resolution better than around four ripples/octave is probably not necessary for the accurate identification of vowels and consonants produced by multiple talkers in quiet backgrounds. The relationship illustrated in Fig. 3 also indicates that spectral peak resolution poorer than around one to two ripples/octave, as seen in many CI listeners and some HI listeners, may result in substantially reduced speech recognition. This finding of an association between severely reduced spectral peak resolution and degraded speech intelligibility in this study is broadly consistent with studies by Baer and Moore (1993), ter Keurs *et al.* (1992), Shannon *et al.* (1995), and others (see the Introduction), who have shown that frequency resolution must be highly impaired in order to severely degrade speech recognition in quiet, and, further, it provides a quantification of the limits of spectral peak resolution below which significant degradation in speech recognition may occur.

While the stimulus used in the spectral peak resolution task was broadband, accurate performance did not require broadband analysis of the signal. Listeners may have used a specific region in which their ability to resolve spectral peaks was particularly good. The specific region of the frequency band the listeners were using for their discrimination was not determined in this study. In addition, listeners may have potentially used level cues at the lower or upper spectral edges of the rippled noise stimuli, since the stimuli were not tapered at the spectral edges (apart from the speech spectrum shaping), and the random variation in the presentation level of the stimuli within an 8 dB range (see Sec. II) may not have been sufficient to eliminate these cues. Despite these limitations, the task provided a strong prediction of speech recognition, with some implant listeners behaving as if they were limited essentially to a single channel, and normal-hearing listeners showing much finer spectral resolution. Further research is required to determine which frequency region(s) are used by individual listeners to perform the task, and the potential perception of level cues at the spectral edges of the stimuli.

These results may have important implications for speech recognition in both HI and CI listeners. Current CI devices and speech processing strategies preserve only crude spectral information. Indeed, while spectral peak resolution varied among CI listeners, and some CI listeners showed performance in the range of the better HI subjects, spectral resolution was poorest in CI listeners on average (see Fig. 2). In addition, while many HI listeners showed spectral peak resolution within the normal range, some HI listeners showed substantially reduced spectral peak resolution. These results indicate that efforts to improve spectral resolution in HI listeners via improved hearing aids, and in CI listeners via improved electrode arrays and speech processing strategies, may result in improved speech recognition.

Turning to the relationships between spectral peak resolution and speech recognition for the individual clinical groups, there was a moderate correlation between spectral peak resolution and both vowel and consonant recognition within the CI listener group, and a fairly weak but significant correlation between spectral peak resolution and vowel and consonant recognition within the HI listener group. The spectral peak resolution test has potential clinical applications in the prediction of speech recognition in HI and CI listeners, as well as in optimizing speech recognition in CI listeners by using the test to determine in a time-efficient manner which speech processing strategy and particular speech processing parameters may provide the best spectral peak resolution for an individual listener. Further research is required to determine whether optimization of the spectral peak resolution test, for instance, by measuring spectral peak resolution in different frequency regions, may improve the predictive value of this test. For CI listeners, while a variety of different speech processing strategies and parameter settings were included in this study in order to specifically assess the relationship between spectral peak resolution and speech recognition with each subject's everyday map, further research is required to examine the effects of these strategies and parameters on both spectral peak resolution and speech recognition. Finally, it seems likely that perceptual factors in addition to spectral peak resolution may contribute to deficits in speech recognition. For example, in cases of poor spectral peak resolution, listeners may rely more on temporal aspects of the speech signal. Modeling other perceptual factors, such as temporal resolution together with spectral peak resolution, may account for a higher amount of variance in speech recognition. Such modeling should explore the inclusion of audibility measures for HI listeners, since audibility is a primary factor related to performance variability in these listeners.

It is important to consider the fact that the findings reported in the present study apply to speech recognition in quiet backgrounds. It is well known, however, that speech recognition in HI and CI listeners is highly susceptible to the effects of competing backgrounds, and it is likely that this is due to reduced spectral resolution. Research suggests that spectral smearing has a more detrimental effect on speech recognition in NH listeners when the processed speech signal is presented in competing backgrounds compared to quiet listening conditions (Baer and Moore, 1993, 1994; Boothroyd *et al.*, 1996; ter Keurs *et al.*, 1992, 1993). In addition, CI simulations in NH listeners indicate that a higher number of spectral channels is required when listening in competing backgrounds compared to quiet listening conditions. Friesen *et al.* (2001) showed an increase in performance in competing backgrounds as the number of channels was increased to 20 channels, which was the highest number tested. Importantly, however, CI listeners in that study did not show a similar increase in performance as the number of channels was increased to 20, but rather showed an asymptote in performance on average with between 4 and 7 channels (depending on the speech material). Further research is required to quantify the role of reduced spectral peak resolution in speech recognition in competing backgrounds.

V. CONCLUSIONS

A direct method of measuring the ability to resolve spectral peaks in complex acoustic spectra was applied to NH, HI, and CI listeners in this study. This method enabled a comparison of spectral peak resolution between NH, HI, and CI listener groups, and an investigation of the relationship between spectral peak resolution and speech recognition across a wide range of perceptual abilities.

The principal findings were as follows.

(1) Spectral peak resolution varied widely among listeners, from 0.13 to 7.55 ripples/octave. The average spectral peak resolution was 4.84 ripples/octave in NH listeners (2.03–7.55 ripples/octave), 1.77 ripples/octave in HI listeners (0.33–4.97 ripples/octave), and 0.62 ripples/octave in CI listeners (0.13–1.66 ripples/octave).

(2) There was a significant relationship between spectral peak resolution and both vowel and consonant recognition across the NH, HI, and CI listener groups, suggesting that the ability to resolve spectral peaks in a complex acoustic spectrum is associated with accurate speech recognition.

(3) There was a plateau in vowel and consonant recognition at around four ripples/octave, indicating that spectral peak resolution better than around four ripples/octave is probably not necessary for the accurate identification of vowels and consonants produced by multiple talkers in quiet backgrounds.

(4) Both vowel and consonant recognition performance deteriorated rapidly when spectral peak resolution fell below one to two ripples/octave.

(5) There was a significant but quite weak linear correlation between spectral peak resolution thresholds and both vowel and consonant recognition for the HI listeners and a significant moderate linear correlation for the CI listeners.

ACKNOWLEDGMENTS

We are grateful to the subjects for their considerable time and effort in participating in this research. We wish to thank Ken Grant, Brian Moore, and Monita Chatterjee for their very helpful comments on previous versions of this paper. Thanks also to Norbert Dillier and Waikong Lai for use of the SCILAB program, and also to Arik Wald and Stephanie Schultz for their assistance. Funding for this research was provided in part by Grant No. 1RO1DC000377 from the National Institutes on Deafness and Other Communicative Disorders, National Institutes of Health and Grant No. RR00059 from the General Clinical Research Centers Program NCRR, National Institutes of Health.

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.

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

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

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Kotby, M. N., Nasser, N. H. A., El Kholy, W. A. H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G. I., Westerman, S., and Ludvigsen, C. (**1994**). "An international comparison of long-term average speech spectra," J. Acoust. Soc. Am. **96**, 2108–2120.

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

Collins, L. M., Zwolan, T. A., and Wakefield, G. H. (**1997**). "Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects," J. Acoust. Soc. Am. **101**, 440–455.

Dallos, P., Ryan, A., Harris, D., McGee, T., and Ozdamar, O. (**1977**). "Cochlear frequency selectivity in the presence of hair cell damage," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic Press, New York), pp. 249–258.

Donaldson, G. S., and Nelson, D. A. (**2000**). "Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies," J. Acoust. Soc. Am. **107**, 1645–1658.

Dorman, M. F., and Loizou, P. C. (**1997**). "Speech intelligibility as a function of the number of channels of stimulation for normal-hearing listeners and patients with cochlear implants," Am. J. Otolaryngol. **18**, S113–S114.

Dorman, M. F., Loizou, P. C., and Rainey, D. (**1997**). "Speech understanding as a function of the number of channels of stimulation for processors using sine-wave and noise-band outputs," J. Acoust. Soc. Am. **102**, 2403–2411.

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

Dreschler, W. A., and Plomp, R. (**1985**). "Relations between psychophysical data and speech perception for hearing-impaired subjects. II," J. Acoust. Soc. Am. **78**, 1261–1270.

Dubno, J. R., and Dirks, D. D. (**1989**). "Auditory filter characteristics and consonant recognition for hearing-impaired listeners," J. Acoust. Soc. Am. **85**, 1666–1675.

Dubno, J. R., Dirks, D. D., and Ellison, D. E. (**1989**). "Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. II: Articulation index predictions," J. Acoust. Soc. Am. **85**, 355–364.

Fay, R. R., Yost, W. A., and Coombs, S. (**1983**). "Psychophysics and neurophysiology of repetition noise processing in a vertebrate auditory system," Hear. Res. **12**, 31–55.

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

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

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

Glasberg, B., 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. (**1989**). "Psychoacoustic abilities of subjects with unilateral and bilateral cochlear impairments and their relationship to the ability to understand speech," Scand. Audiol. Suppl. **32**, 1–25.

Henry, B. A., and Turner, C. W. (**2003**). "The resolution of complex spectral patterns in cochlear implant and normal hearing listeners," J. Acoust. Soc. Am. **113**, 2861–2873.

Henry, B. A., McKay, C. M., McDermott, H. J., and Clark, G. M. (**2000**). "The relationship between speech perception and electrode discrimination in cochlear implantees," J. Acoust. Soc. Am. **108**, 1269–1280.

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

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

Humes, L. E., Dirks, D. D., Bell, T. S., Ahlstrom, C., and Kincaid, G. E. (**1986**). "Application of the articulation index and the speech transmission index to the recognition of speech by normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **29**, 447–462.

Lai, W. K., Bogli, H., and Dillier, N. (**2003**). "A software tool for analyzing multichannel cochlear implant signals," Ear Hear. **24**, 380–391.

Levitt, H. (**1971**). “Transformed up-down methods in psychoacoustics,” J. Acoust. Soc. Am. **49**, 467–477.

Liberman, M. C., and Dodds, L. W. (**1984**). “Single-neuron labeling and chronic cochlear pathology: III. Stereocilia damage and alterations of threshold tuning curves,” Hear. Res. **16**, 55–74.

Lutman, M. E., and Clark, J. (**1986**). “Speech identification under simulated hearing-aid frequency response characteristics in relation to sensitivity, frequency resolution, and temporal resolution,” J. Acoust. Soc. Am. **80**, 1030–1040.

Nelson, D. A., Van Tassell, D. J., Schroder, A. C., Soli, S., and Levine, S. (**1995**). “Electrode ranking of ‘place pitch’ and speech recognition in electrical hearing,” J. Acoust. Soc. Am. **98**, 1987–1999.

Pavlovic, C. V. (**1984**). “Use of the articulation index for assessing residual auditory function in listeners with sensorineural hearing impairment,” J. Acoust. Soc. Am. **75**, 1253–1258.

Pavlovic, C. V., Studebaker, G. A., and Sherbecoe, R. L. (**1986**). “An articulation index based procedure for predicting the speech recognition performance of hearing-impaired individuals,” J. Acoust. Soc. Am. **80**, 50–57.

Seligman, P. M., and McDermott, H. J. (**1995**). “Architecture of the Spectra 22 speech processor,” Ann. Otol. Rhinol. Laryngol. Suppl. 104, 139–141.

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

Skinner, M. W. (**1980**). “Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation,” J. Acoust. Soc. Am. **67**, 306–317.

Skinner, M. W., Arndt, P. L., and Staller, S. J. (**2002**). “Nucleus 24 advanced encoder conversion study: Performance versus preference,” Ear Hear. **23**, 2S–16S.

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.

Supin, A., Popov, V. V., Milekhina, O. N., and Tarakanov, M. B. (**1994**). “Frequency resolving power measured by rippled noise,” Hear. Res. **78**, 31–40.

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**). “Limited resolution of spectral contrast and hearing loss for speech in noise,” J. Acoust. Soc. Am. **94**, 1307–1314.

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.

Trees, D. A., and Turner, C. W. (**1986**). “Spread of masking in normal subjects and in subject with hearing loss,” Audiology **25**, 70–83.

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

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.

Tyler, R. S., Wood, E. J., and Fernandez, M. (**1982**). “Frequency resolution and hearing loss,” Br. J. Audiol. **16**, 45–83.

Vandali, A. E., Whitford, L. A., Plant, K. L., and Clark, G. M. (**2000**). “Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system,” Ear Hear. **21**, 608–624.

Wightman, F. L., McGee, T., and Kramer, M. (**1977**). “Factors influencing frequency selectivity in normal and hearing-impaired listeners,” in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 295–306.

Wilson, B. S., Finley, C. F., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (**1991**). “Better speech recognition with cochlear implants,” Nature (London) **352**, 236–238.

Yost, W. A., Patterson, R., and Sheft, S. (**1996**). “A time domain description for the pitch strength of iterated rippled noise,” J. Acoust. Soc. Am. **99**, 1066–1078.

Zwolan, T. A., Collins, L. M., and Wakefield, G. H. (**1997**). “Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects,” J. Acoust. Soc. Am. **102**, 3673–3685.