

Psychoacoustic and Phoneme Identification Measures in Cochlear-Implant and Normal-Hearing Listeners

Trends in Amplification
17(1) 27–44
© The Author(s) 2013
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1084713813477244
tia.sagepub.com


Ray L. Goldsworthy, PhD¹, Lorraine A. Delhorne, MS¹,
Louis D. Braida, MS, PhD¹, and Charlotte M. Reed, PhD¹

Abstract

The purpose of this study is to identify precise and repeatable measures for assessing cochlear-implant (CI) hearing. The study presents psychoacoustic and phoneme identification measures in CI and normal-hearing (NH) listeners, with correlations between measures examined. Psychoacoustic measures included pitch discrimination tasks using pure tones, harmonic complexes, and tone pips; intensity perception tasks included intensity discrimination for tones and modulation detection; spectral-temporal masking tasks included gap detection, forward and backward masking, tone-on-tone masking, synthetic formant-on-formant masking, and tone in noise detection. Phoneme perception measures included vowel and consonant identification in quiet and stationary and temporally gated speech-shaped noise. Results on psychoacoustic measures illustrate the effects of broader filtering in CI hearing contributing to reduced pitch perception and increased spectral masking. Results on consonant and vowel identification measures illustrate a wide range in performance across CI listeners. They also provide further evidence that CI listeners obtain little to no release of masking in temporally gated noise compared to stationary noise. The forward and backward-masking measures had the highest correlation with the phoneme identification measures for CI listeners. No significant correlations between speech reception and psychoacoustic measures were observed for NH listeners. The superior NH performance on measures of phoneme identification, especially in the presence of background noise, is a key difference between groups.

Keywords

cochlear implants, speech reception, psychoacoustics

Introduction

This study examines relationships between psychoacoustic measures using stimuli consisting of tones and noises with measures of consonant and vowel identification in quiet, speech-shaped noise and gated speech-shaped noise in normal-hearing (NH) and cochlear-implant (CI) listeners. The purpose is to identify psychoacoustic measures that have clinical relevance to speech perception in cochlear implantees. To this end, this study assesses a wide range of psychoacoustic measures that are relatively efficient to implement.

Previous studies have established some relationships between speech reception and basic psychoacoustic abilities in NH and hearing-impaired listeners. Studies with NH listeners, as defined by audiometric loss, generally have found little to no correlation between speech reception and psychoacoustic abilities (Strouse et al., 1998; Kidd, Watson, & Gygi, 2007; Surprenant & Watson, 2001; Watson & Kidd, 2002). A general conclusion provided by Kidd et al. is that differences between NH listeners on speech reception tasks are better attributed to differences in familiar sound recognition as

opposed to differences in psychoacoustic abilities. When considering hearing-impaired listeners, more evidence has been found linking speech reception with psychoacoustic abilities. The strongest correlation between speech reception and an underlying psychoacoustic ability is generally found with audiometric loss (e.g., Dreschler & Plomp, 1980; Dubno & Schafer, 1992). But studies have also shown correlations when considering measures beyond audibility. For example, there have been several studies that have demonstrated correlations between speech reception and measures of temporal and spectral processing (Boothroyd, Mulhearn, Gong, & Ostroff, 2001; Dreschler & Plomp, 1985; Festen & Plomp, 1983; Thibodeau & Van Tasell, 1987; van Rooij & Plomp, 1990; van Rooij, Plomp, & Orlebeke, 1989; van Schijndel,

¹Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA

Corresponding Author:

Ray L. Goldsworthy, Massachusetts Institute of Technology, Building 36
Room 747, 77 Massachusetts Ave., Cambridge, MA 02139, USA.
Email: raygold@sens.com

Houtgast, & Festen, 2001a, 2001b; Kishon-Rabin et al., 2009). Festen and Plomp, for example, concluded that speech reception in noise was correlated to spectral processing but that speech reception in quiet is governed by audiometric loss.

Previous research has also been directed toward understanding the speech reception abilities of individual listeners with cochlear implants in terms of their performance on basic psychoacoustic tasks. Several studies have examined the relationship between speech reception through a cochlear implant and single-electrode psychoacoustic measures. Previous investigators have concluded that gap detection (Busby & Clark, 1999; Shannon, 1989), forward masking recovery time (Shannon, 1990), and electrode discrimination (Zwolan, Collins, & Wakefield, 1997) are poor predictors of speech reception. However, Busby and Clark did find a correlation with gap detection when examining auditory-visual integration in an aided speech-reading task.

In contrast to the aforementioned studies, Fu and colleagues have observed various correlations between CI speech reception and measures of psychoacoustic ability. Using vocoder simulations of CI processing with varying number of spectral channels, Fu, Shannon, and Wang (1998) demonstrated that speech reception in noise strongly depends on spectral resolution even though speech reception in quiet is only marginally dependent. Fu (2002) found a high correlation between speech reception and a single-electrode temporal modulation detection task. He postulated that he was able to detect this correlation by examining modulation thresholds over a range of loudness values and by restricting listeners to have similar sound processing strategies. Correlations were also found by Pfungst and Xu (2005) who reported on relationships between clinical psychoacoustic measures with speech reception scores in quiet. They found that implant users with high variations in electrical stimulation comfort levels (across stimulating electrodes) had poorer speech recognition. They also found that subjects with high mean comfort levels and large mean dynamic ranges had better speech recognition.

The use of a spectrally rippled stimuli processed through an implant user's clinical processor has had relatively wide success as a predictor of speech reception in both quiet and noise (Supin et al., 1994; Henry & Turner, 2003; Henry, Turner, & Behrens, 2005; Litvak, Spahr, Saoji, & Fridman, 2007; Won, Drennan, & Rubinstein, 2007). Won et al. found that test-retest reliability for spectral-ripple discrimination to be good and observed no learning for listeners on the measure. They therefore concluded that spectral-ripple discrimination would serve as a reliable tool to evaluate CI performance under different signal processing strategies.

In summary, although there is agreement that speech reception in quiet is primarily governed by audiometric loss, there is less agreement regarding how psychoacoustic measures are related to speech reception in noise, although many of the aforementioned studies have concluded that spectral resolution is an underlying determinant. The purpose of the study presented here is to further examine the relationships between psychoacoustic measures and phoneme perception

in NH and CI listeners. Establishing such correlations is of value not only for predicting potential speech reception results of CI listeners but also as a guide for developing improved signal-processing schemes for cochlear implants.

Method

Listeners

The listeners for this study included 10 adults with NH and 12 adult and 1 adolescent CI listeners. All listeners provided informed consent on their first visit to the laboratory and were paid for their participation in the study.

NH listeners. Listeners ranged aged between 19 and 65 years at the time of testing ($M = 45$ years; male, $n = 2$; female, $n = 8$). A hearing test was administered to listeners before they began the experiment to document that they had clinically normal hearing (defined as 20-dB HL or better at the octave frequencies between 125 and 8000 Hz). These listeners had all participated in other studies in our laboratory and were originally recruited through ads posted in the local community.

CI listeners. Demographic and audiologic information about each CI listener is listed in Table 1. The listeners ranged in age between 16 and 66 years at the time of testing ($M = 52$ years; male, $n = 6$; female, $n = 7$). Five of the listeners (CI_3 , CI_5 , CI_7 , CI_{10} , and CI_{11}) had binaural implants but only CI_3 , CI_{10} , and CI_{11} had sufficient participation to separately measure each of their ears. Thus, data are reported for 16 ears arising from 13 listeners. Eleven of the 13 listeners reported that the cause of their hearing loss was either a genetic disorder (CI_1 , CI_2 , CI_3 , CI_9 , and CI_{15}) or was unknown (CI_6 , CI_7 , CI_{10} , CI_{11} , CI_{12} , and CI_{13}). In these cases, the loss was typically diagnosed in early childhood and progressed over time. Listener CI_5 had normal hearing until age 6 when she contracted mumps, which resulted in a hearing loss that progressed over time. Listener CI_{18} had normal hearing until he was 32 and lost his hearing to ototoxicity. The age at implantation in each of the 16 test ears ranged from 15 to 65 ($M = 45$) years and duration of implant use ranged from 1 to 25 years. Two listeners had worn their implants for more than 20 years, 4 listeners between 5 and 8 years, and 10 listeners for less than 5 years. The listeners used a variety of CI sound processors, including the Auria (three ears), Harmony (five ears), Nucleus 3G (three ears), Nucleus Freedom (two ears), Nucleus N5 (one ear), and Geneva (two ears) devices.

None of the CI listeners had prior experience with psychoacoustic or speech testing aside from the procedures used in standard clinical visits for sound processor mapping and assessment.

Psychoacoustic Measures

Psychoacoustic measures were implemented using an adaptive forced-choice paradigm. (The letter code used in Table 2 to identify each of the measures is provided below.) The six

Table 1. Demographics of the Cochlear Implantees.

Listener	Ear	Etiology	Age at onset of hearing loss/ deafness	Age at implantation (years)	Implant use (years)	Implant processor
CI1	R	Genetic disorder	Birth progressive	63	3	AB/AURIA
CI2	R	Secondary to inherited syndrome	Birth progressive	21	2	AB /Harmony
CI3	L	Genetic disorder	Birth progressive	60	6	Cochlear/Esprit 3G
CI3	R	Genetic disorder	Birth progressive	60	7	Cochlear/Esprit 3G
CI5	R	Mumps	Diagnosed at age 6, progressive	50	2	AB /Harmony
CI6	L	Unknown.	Diagnosed at age 27, progressive	56	7	Cochlear/Esprit 3G
CI7	R	Unknown	Diagnosed at age 3.5, progressive	59	25	AB/Geneva
CI9	R	Genetic	Diagnosed at age 8, progressive	54	3	AB /Harmony
CI10	L	Unknown	Birth progressive	18	2	Cochlear/Freedom
CI10	R	Unknown	Birth progressive	16	1	Cochlear/Freedom
CI11	L	Unknown	Diagnosed at age 5, progressive	63	23	AB/Geneva
CI11	R	Unknown	Diagnosed at age 10, progressive	63	3	AB/AURIA
CI12	L	Unknown	Birth progressive	54	4	AB/AURIA
CI13	L	Unknown	Diagnosed at age 5, progressive	58	2	AB /Harmony
CI15	L	Genetic	Birth progressive	65	1	AB /Harmony
CI18	L	Ototoxicity (Gentamycin)	32 progressive	55	8	Cochlear/Nucleus 5

Table 2. Correlation Analysis of the Measures for the Cochlear Implantees (See Text for the Coding of Measure).

	PT	PTR	TP	TPR	HC	HCR	PTI	PTIR	MD	G	FM	BM	TiN	ToT	FoF	CQ	CS	CG	CM	VQ	VS	VG	
PTR	.93																						
TP	.82	.70																					
TPR	.89	.88	.84																				
HC	.85	.89	.57	.75																			
HCR	.90	.94	.67	.80	.85																		
PTI	.54	.59	.50	.43	.39	.69																	
PTIR	.02	.13	-.11	.00	.24	.19	.23																
MD	-.09	.01	-.26	-.22	.12	.15	.23	.37															
G	.36	.60	.08	.52	.51	.54	.13	.30	.17														
FM	.47	.54	.44	.63	.53	.45	.30	-.10	.29	.39													
BM	.45	.49	.42	.55	.49	.52	.26	-.16	.36	.40	.87												
TiN	.37	.54	.27	.53	.36	.49	.57	-.08	.18	.47	.67	.67											
ToT	.54	.60	.45	.66	.57	.52	.33	-.23	.13	.34	.86	.81	.78										
FoF	.54	.57	.52	.61	.52	.60	.10	-.08	.15	.49	.67	.70	.37	.46									
CQ	-.34	-.43	-.39	-.52	-.33	-.41	-.08	.04	-.25	-.58	-.69	-.73	-.45	-.45	-.79								
CS	.64	.62	.56	.58	.59	.72	.57	-.09	.13	.52	.23	.42	.30	.14	.51	-.68							
CG	.60	.52	.56	.69	.55	.56	.37	-.24	.15	.22	.88	.93	.80	.87	.73	-.67	.38						
CM	-.41	-.32	-.40	-.52	-.38	-.34	-.19	.22	-.12	-.04	-.86	-.84	-.75	-.92	-.59	.46	-.03	-.94					
VQ	-.44	-.51	-.36	-.56	-.54	-.48	.01	-.08	-.22	-.64	-.64	-.68	-.32	-.47	-.76	.84	-.56	-.65	.49				
VS	.10	.04	.07	.10	-.07	.15	.02	-.01	.50	.27	.12	.15	-.03	-.17	.50	-.59	.44	.17	-.02	-.48			
VG	.57	.48	.55	.59	.34	.56	.47	-.16	.42	.28	.65	.70	.48	.50	.72	-.84	.60	.73	-.56	-.77	.71		
VM	-.70	-.61	-.69	-.73	-.50	-.66	-.62	.21	-.25	-.21	-.79	-.85	-.66	-.77	-.65	.76	-.53	-.87	.74	.73	-.32	-.90	

measures that explicitly quantified pitch perception used a two-interval procedure, and listeners were instructed to judge which interval was higher in pitch. The other nine measures used a three-interval procedure, and listeners were instructed to judge the interval that was different from the other two. The intensity of the stimuli was adjusted to a comfortable level prior to each measure. Listeners were instructed as to

how to compare the sounds. Responses were gathered using a computer graphic user interface. Trial-by-trial correct-answer feedback was provided visually in all testing. If the task required the listener to judge which interval had the higher pitch then trials were scored as correct and feedback given accordingly when the listener correctly identified the target interval containing the higher-frequency sound.

The adaptive rule used for each measure was a 1-up, 2-down rule that converged to a 70.7% probability of correct response. The initial value of the adaptive variable was set such that all listeners could readily discriminate between standard and target sounds. Adaptive measures used either a geometric or a linear rule for controlling the adaptive variable. For the geometric rule, the adaptive variable was reduced (or increased) after each reversal by a constant ratio. This ratio had an initial value of 50% and was itself reduced by a factor of 0.707 after each reversal. For the linear rule, the adaptive variable was reduced (or increased) after each reversal by a specified additive step. This step had an initial value defined for each measure and was itself reduced by a factor of 0.707 after each reversal. The adaptive procedures terminated after six reversals, and the threshold for the adaptive variable was taken as the final value of the adaptive variable. Within a given measure, each condition was tested using three runs in succession, randomized across conditions.

Stimulus intervals were always 400 ms in duration with a 200-ms interstimulus interval. Unless noted otherwise, all intervals were ramped on and off with a logarithmic ramp. These ramps started 64 dB below the signal's peak amplitude and had a constant rate of increase (or decrease) over the corresponding 8-ms attack and decay times. All processing filters had an infinite impulse response. Stimulus intervals were visually cued with buttons on a computer monitor. When noise was used, it was drawn from a uniform distribution and processed through filters as noted for the specific measure.

All stimuli were generated digitally in Matlab[®] on a laboratory computer. The digitized signals were played through a sound card (Lynx Studio Lynx One) with a sampling frequency of 24 kHz and 24-bit resolution. The auditory signal was sent through the sound card to an attenuator TDT PA4 and headphone buffer TDT HB6 before being presented monaurally through headphones (Sennheiser HD 580) to listeners who were seated in a double-walled soundproof room. The measures described below required three, 2-hr, sessions per listener to complete. Listeners were allowed breaks as needed.

Pitch perception. Pure-tone difference limens for frequency: The standard interval was a 400-ms pure tone, whose frequency was defined as the condition variable. Frequencies tested were 500, 1000, and 2000 Hz. The target interval was identical to the standard except that the tone frequency was adaptively higher in frequency than the standard. The initial frequency difference was 100% and was adapted using the geometric rule. Listeners were instructed to select the interval that sounded higher in pitch.

Pure-tone difference limens for frequency with intensity and frequency roving (PTR): This measure is equivalent to the *pure-tone difference limen* measure with the exception that the tones were roved within a $\pm 50\%$ range for frequency and were roved across intervals within ± 6 -dB range for intensity. The frequency of the target tone was adaptively higher than the standard as in the preceding measure. The

listeners were instructed to attend only to the pitch of the stimuli and to select the interval that sounded higher in pitch.

Tone-pip (TP) DLs for frequency: The standard TP was the impulse response of a second-order band-pass filter, the center frequency of which was defined as the condition variable and the bandwidth (3 dB down) of which was 0.29% of the center frequency. Frequencies tested were 500, 1000, and 2000 Hz. The target sound was generated identically as the standard except that the center frequency of the generating filter was adaptively higher than the standard. The starting frequency difference was 100% and was adapted using the geometric rule. The resulting signals were truncated to 400 ms and ramped on and off using 8-ms ramps. The listener's task was to select the interval that sounded higher in pitch.

Tone-pip difference limens for frequency with intensity and frequency roving (TPR): This measure is equivalent to the previous measure with the exception that the tones were roved across trials within a $\pm 50\%$ range for frequency, and across intervals within a ± 6 -dB range for intensity. The frequency of the target tone was adaptively higher than the standard. Starting value of the target and instructions to listeners were as for the previous measure.

Harmonic complex difference limens for frequency (HC): The standard interval contained a pulse train that was filtered through a second-order filter with a 1000-Hz center frequency and a $\frac{1}{2}$ -octave bandwidth. The condition variable was the frequency of the pulse train, and the values tested were 110, 220, and 440 Hz. The target interval was generated identically as the standard except that the center frequency of the generating filter was adaptively higher than the standard. The starting value of the frequency difference was 100% and was adapted using the geometric rule. The listener's task was to select the interval that sounded higher in pitch.

Harmonic complex difference limens for frequency with intensity and frequency roving (HCR): This measure is equivalent to the previous measure with the exception that the tones were roved across trials within a $\pm 50\%$ range for frequency and across intervals within a ± 6 -dB range for intensity. The frequency of the target tone was adaptively higher than the standard. Starting value of the adaptive variable and instructions to listeners were identical to that used in the previous measure.

Intensity perception. Pure-tone difference limens for intensity (PTI): The standard interval contained a 1000-Hz pure tone, the intensity of which was the condition variable. Intensities tested were 0, -12, and -24 dB relative to the listener-selected comfortable listening level. The target interval was identical to the standard except that the stimulus intensity was adaptively higher in intensity. The initial value of the adaptive variable was 12 dB higher than the standard intensity and was adapted using the geometric rule. Listeners were instructed to select the interval that was louder.

Pure-tone difference limens for intensity with intensity and frequency roving (PTIR): This measure is equivalent to the previous measure with the exception that the tones were

roved across intervals within a $\pm 50\%$ range for frequency and across trials within a ± 6 -dB range for intensity. Starting value of the adaptive variable and instructions to listeners were as for the previous measure.

Difference limens for modulation-depth detection of a modulated pure tone (MD): The standard interval contained a 1000-Hz pure tone. The target was as the standard but with a multiplicative sinusoidal modulator, the frequency of which was the condition variable taking values of 10 and 100 Hz. The modulation depth was the adaptive variable with an initial value of 100% following the geometric rule. Listeners were given examples of the measurement stimuli, instructed on how to attend to the target, and instructed to select the corresponding interval.

Spectral-temporal masking. Gap detection in filtered noise (G): The standard interval was generated by filtering random noise through a second-order Butterworth filter with a 1000-Hz center frequency and a 414-Hz bandwidth. The target was generated as the standard except there was a gap of silence commencing 200 ms into the stimulus. The gap was gated using 4-ms fall and rise times. Gap duration was the adaptive variable with an initial value of 80 ms following the geometric rule. Listeners were instructed on how to listen for the gap and to select the interval with the stimulus containing the gap.

Forward masking of tones on tonal targets (FM): In the forward-masking measure, the standard was a 300-ms, 1000-Hz, pure tone followed by 100 ms of silence to complete the 400-ms interval. The target was as the standard but with a 10-ms tone pulse immediately following the 300-ms tone. Both this 10-ms tone pulse and the 300-ms tone were ramped on and off using 4-ms rise and fall times. The frequency of the 10-ms tone pulse was the condition variable and took values of 600, 800, 1200, and 1600 Hz. The amplitude of the 10-ms tone pulse was the adaptive variable the initial value of which was 12 dB higher than the 300-ms tone. The initial step size was 6 dB and was modified using the linear rule. Listeners were given examples of the measurement stimulus, instructed on how to attend for the brief secondary tone and instructed to select the corresponding interval. Results are reported as the target to masker ratio in dB.

Backward masking of tones on tonal targets (BM): This measure was defined similarly to the previous measure but having the 10-ms target tone pulse precede the 300-ms tone. Starting value of the adaptive variable and instructions to listeners were as for the previous measure.

Tone detection with a secondary tone masker (ToT): The standard interval was a 1000-Hz pure tone. The target was as the standard plus an additional tone the frequency of which was the condition variable and took the values 600, 800, 1200, and 1400 Hz. The intensity of this secondary tone was the adaptive variable and took an initial value of 12 dB higher than the standard tone. The initial step size was 6 dB and was modified using the linear rule. Listeners were instructed to select the interval that sounded different.

Tone detection with a stationary noise masker (TiN): The standard interval was generated by filtering random noise through a second-order Butterworth filter with a 1000-Hz center frequency and with a variable bandwidth. This filtered noise was normalized to have a constant output power as the bandwidth was varied. The target was as the standard plus a 1000-Hz pure tone. The intensity of this secondary tone was the adaptive variable and took an initial value of 12 dB higher than the noise. The initial step size was 6 dB and was modified using the linear rule. Listeners were instructed to select the interval that sounded different. Results are reported as the signal to noise ratio in dB.

Detection of a secondary formant in a harmonic complex (FoF): The standard interval was a 110-Hz pulse train filtered through a second-order Butterworth filter with a 1000-Hz center frequency and a 414-Hz bandwidth. The target interval was as the standard plus a second component that was also a 110-Hz pulse train filtered through a second-order Butterworth filter with a 41.4% bandwidth, but the center frequency of which was the condition variable taking values of 500 and 2000 Hz. The intensity of the secondary filtered pulse train was the adaptive variable and took an initial value of 12 dB higher than the standard interval. The initial step size was 6 dB and was modified using the linear rule. Listeners were instructed to select the interval that sounded different from the other two.

Phoneme Identification

Six measures of phoneme identification were tested that considered vowel and consonant reception in quiet, in stationary noise, and in temporally gated noise. The vowel and consonant paradigm used was as in Fu (Fu & Nogaki, 2005). Vowels were drawn from speech samples collected by Hillenbrand et al. (Hillenbrand, Getty, Clark, & Wheeler, 1995) for five male and five female talkers. The vowel set consisted of ten monophthongs (/i I ε æ u ō a ɔ Λ ɜ/) and two diphthongs (/əʊ eI/), presented in /h/-V-/d/ context (heed, hid, head, had, who'd, hood, hod, hud, hawed, heard, hoed, hayed). Consonants were drawn from speech samples collected by Shannon et al. (Shannon, Jensvold, Padilla, Robert, & Wang, 1999) for five male and five female talkers. The consonants consisted of twenty phonemes /b d g p t k m n l r y w f s ʃ v z ð t ʃ d ʒ/, presented in /a/-C-/a/ context. Each speaker contributed one token of each vowel syllable (resulting in 120 tokens in the vowel set) and one token of each consonant syllable (resulting in 200 tokens in the consonant set). For each test, 20 tokens were selected randomly and used in a 20-item training procedure just prior to testing of the complete corpus of tokens. Stimuli were selected randomly without replacement in all testing. Listeners responded using a 12 (vowels) or 20 (consonants) button graphical user interface with the appropriately labeled phonemes. Feedback was provided in the form of a visual display of "correct" or "wrong" following each trial.

For each speech measure, a letter code is provided for use in Table 2.

Vowel (VQ) and consonant (CQ) identification in quiet. For vowels, the experiment employed a one-interval, 12-alternative forced-choice task (where chance performance was 8.3% correct) with a total of 120 trials (one presentation of each of the vowel tokens). For consonants, the experiment employed a 20-alternative forced-choice task (where chance performance was 5% correct) with a total of 200 trials (one presentation of each of the consonant tokens). On each trial, listeners were instructed to select the alternative that most closely corresponded to the speech syllable that had been presented. The results of the vowel and consonant identification measures are presented in terms of percent correct in rational arcsine units (rau) to equalize variance. Prior to the measure, the listener adjusted the level of a phoneme token to a comfortable level and all subsequent tokens were presented with the same average power.

Vowel and consonants identification in noise. Measures were conducted to determine the noise level at which listeners correctly identified 50% correct of the tokens correctly. Listeners were tested using both stationary speech-shaped noise (VS and CS) and temporally gated speech-shaped noise (VG and CG, henceforth gated noise). The masking-level difference for vowels (VM) and consonants (CM) is the difference between the measure in stationary and gated noise. Speech-shaped noise maskers were generated by passing random noise drawn from a uniform distribution through a speech-shaping filter. This filter was generated by estimating the power spectral density of the corresponding speech corpus using Welch's periodogram and converting this density to an eighth-order IIR filter using Prony's method (e.g., Parks & Burrus, 1987). For the measures using gated noise, the noise source was gated on and off at 10 Hz with 4-ms fall and rise times.

Measures in noise used an adaptive rule to determine the speech to noise ratio (SNR) at which listeners correctly identified 50% correct of the phonemes. Prior to the measure, the listener adjusted the level of a phoneme token to a comfortable level and all subsequent tokens were presented with the same average power. The level of the noise was adapted using a 1-up, 1-down rule that converges to a 50%-correct probability of correct response. The signal to noise ratio (SNR) was the adaptive variable and had an initial value of 12 dB and the step size for adjustment was 6 dB for the first six reversals, 3 dB for the next six, 1.5 dB for the next six, and then 0.75 dB for the remainder of the test. The SNR was taken as the average over the 120 trials of the vowel measure or the 200 trials of the consonant measure (and did not include the 20 training trials).

Not all listeners could achieve this 50%-correct identification criterion. Listeners who could not achieve this criterion (CI_{3L} , CI_{3R} , CI_{5R} , CI_{9R} , CI_{11R} , CI_{12L} , and CI_{15L}) were tested on modified phoneme identification in noise measures. For example, they were tested using smaller phoneme sets or using a lower identification criterion. Such

modifications to the measurement protocol make it difficult to conduct fair comparisons across listeners; consequently, we only consider listeners who perform the measure using the standard 50%-correct rule with the full phoneme sets when conducting subsequent correlation analysis. Listeners (and respective ear) included in this analysis were: CI_{1R} , CI_{2R} , CI_{6L} , CI_{7R} , CI_{10L} , CI_{10R} , CI_{11L} , CI_{13L} , and CI_{18L} .

Results

For each of the psychoacoustic and speech reception measures, results are presented in Appendices 1 through 4. Individual results from the 13 cochlear implantees and average results and ranges from the 9 NH listeners are given. For each measure ANOVAs were conducted to investigate the effects of stimulus parameters and listener group.

The results for the CH and NH listeners were also subjected to correlation analyses. The correlation coefficient between pairs of measurements was computed for each pair of tests (averaged across conditions) and for both listener groups. Averaging across conditions allowed for a single value for each measure to be considered. This averaging was partially justified since condition generally was not a significant factor in the analysis. With such a large number of correlations considered, we used a relatively stringent criterion for significance. In particular, we report correlations as significant when the family-wise error rate is $p < .05$ and as moderately significant when the family-wise error rate is $.05 < p < .1$. These family-wise error rates are adjusted using the Bonferroni correction, which simply divides these rates by the number of measures under consideration resulting in error rates of $p < .0024$ and $p < .0048$. The results of the correlation analyses for the cochlear implantees are reported in Table 2. Although correlation results from listeners with normal hearing will be mentioned below, they generally did not reach significance and thus are not reported in detail.

Psychoacoustic Measures

Pitch perception. Appendix 1 and Figure 1 show performance for NH and CI listeners. Both groups of listeners showed a wide range of performance on the six pitch perception measures with difference limens generally in the range of 5% to 30%. These limens are much worse than published results that use listeners selected as having musical experience and/or are trained specifically on the pitch perception task (Kidd et al., 2007; Micheyl, Delhommeau, Perrot, & Oxenham, 2006). Observed FDLs on our task spanned a wide range with the best/worst performing NH listeners scoring 0.2/30% and the best-/worst-performing CI listeners scoring 0.5/89%.

Figure 1 shows the effect of roving on frequency discrimination for pure tones, harmonic complexes, and tone pips. For all three stimuli, the effect of roving is nearly the same for the CI and NH listeners: Roved difference limens are larger than nonroved limens. The effect is larger for CI than

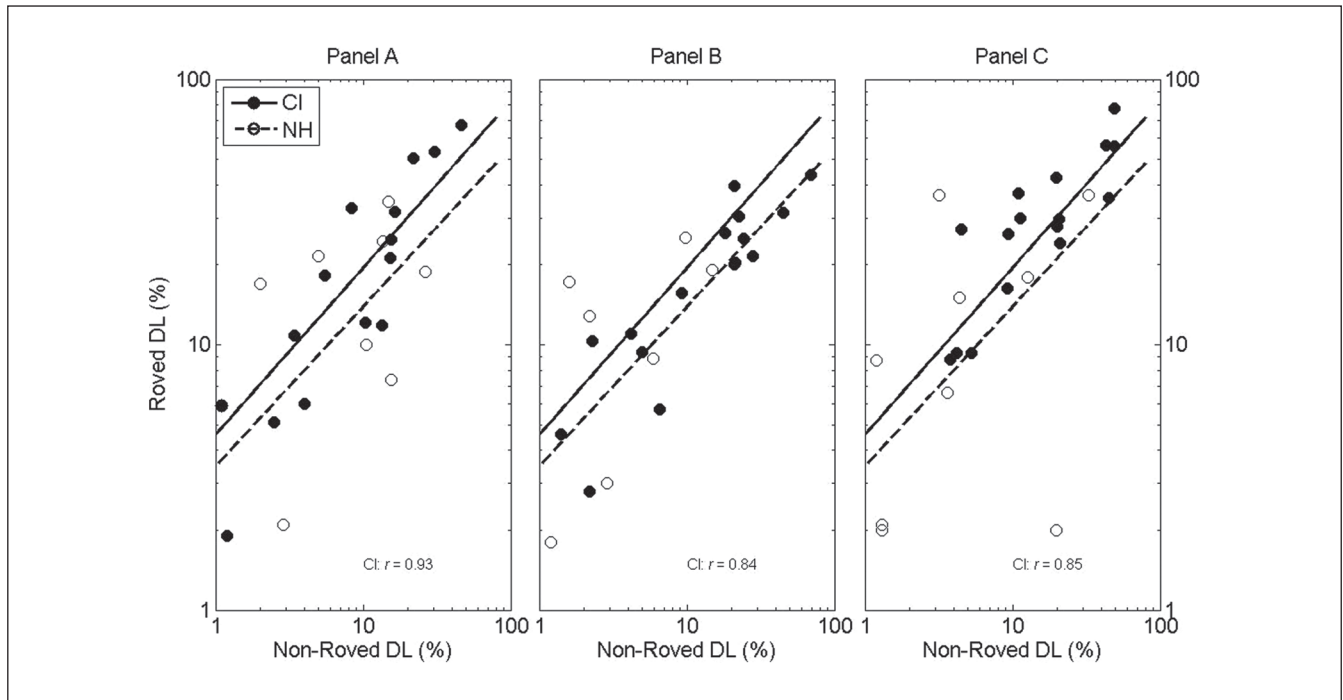


Figure 1. Roved versus nonroved difference limens for pure tones (Panel A), harmonic complexes (Panel B), and tone pips (Panel C).

for NH listeners and relatively larger for listeners with smaller limens.

An analysis of variance was computed on the pitch perception data using group (i.e., CI or NH), measure (i.e., pure tone, tone pips, or harmonic complex), roving, and condition as factors. Group was significant, $F(1, 432) = 84.8, p < .001$, with an NH average of 4.8% and a CI average of 14.2%. Roving was significant, $F(1, 432) = 36.9, p < .001$, with roving increasing average FDLs from 6.6 to 12.9%. Neither measure, $F(2, 432) = 1.9, p = .16$, nor condition, $F(2, 432) = .11, p = .90$, was significant. The interaction between group and measure was significant, $F(2, 432) = 5.5, p = .004$, with CI listeners tending to perform better on the pure-tone measure (10.3%) compared to the tone-pip (13.6%) and harmonic complex (18.9%) measures. The NH averages were 5.9, 3.7, and 4.9% for the pure-tone, tone-pip, and harmonic complex measures, respectively. No other interactions between factors were significant at $p < .1$.

Intensity perception. Analysis of variance was computed on the pure-tone difference limens for intensity shown in Appendix 2 using group, roving, and condition as factors. Group was not significant, $F(1, 144) = .6, p = .14$. Roving was significant, $F(1, 144) = 10.2, p < .001$, increasing average IDLs from 2.7 to 4.5 dB. Condition was not significant, $F(2, 144) = 0.6, p = .55$. No interactions between factors were significant at $p < .2$.

Analysis of variance was computed on the difference limens for modulation detection using group and condition as factors. Group was significant, $F(1, 48) = 7.9, p = .007$, with an NH average of 16.1% and a CI average of 28.3%.

Neither condition, $F(1,48) = .06, p = .81$, nor the interaction between group and condition, $F(1, 48) = 1.1, p = .30$, were significant.

Spectral-temporal masking. Data from the six spectral-temporal masking tasks are presented in Appendix 3.

The results of an analysis of variance conducted on the gap detection thresholds using group as a factor indicated that group was moderately significant, $F(1, 22) = 5.0, p = .036$, with an NH average of 8.1 ms and a CI average of 6.2 ms.

Figure 2 plots the forward- and backward-masking thresholds for different target frequencies. Clearly the relationship between these two types of masking is very similar for CI and NH listeners, although CI listeners exhibit greater amounts of masking than do NH listeners. Analysis of variance was computed on these thresholds using group measure (i.e., backward or forward masking) and condition as factors. Group was significant, $F(1, 180) = 27.7, p < .001$, with average thresholds for NH and CI listeners of -15.8 and -7.0 dB, respectively. Measure was not significant, $F(1, 180) = 1.2, p = .27$. Condition was significant, $F(3, 180) = 5.5, p = .0012$. No interactions between factors were significant at $p < .2$.

Analysis of variance was computed on the tone in noise thresholds using group and condition as factors. Group was significant, $F(1, 92) = 50.5, p < .001$, with an average NH threshold of -6.1 dB and an average CI threshold of 3.3 dB. Both condition, $F(3, 92) = 7.7, p < .001$, and the interaction between group and condition, $F(3, 92) = 4.0, p = .01$, were significant. These effects can be observed in the plot of tone in noise discrimination thresholds versus noise bandwidth given in Figure 3. Discrimination thresholds improve as the

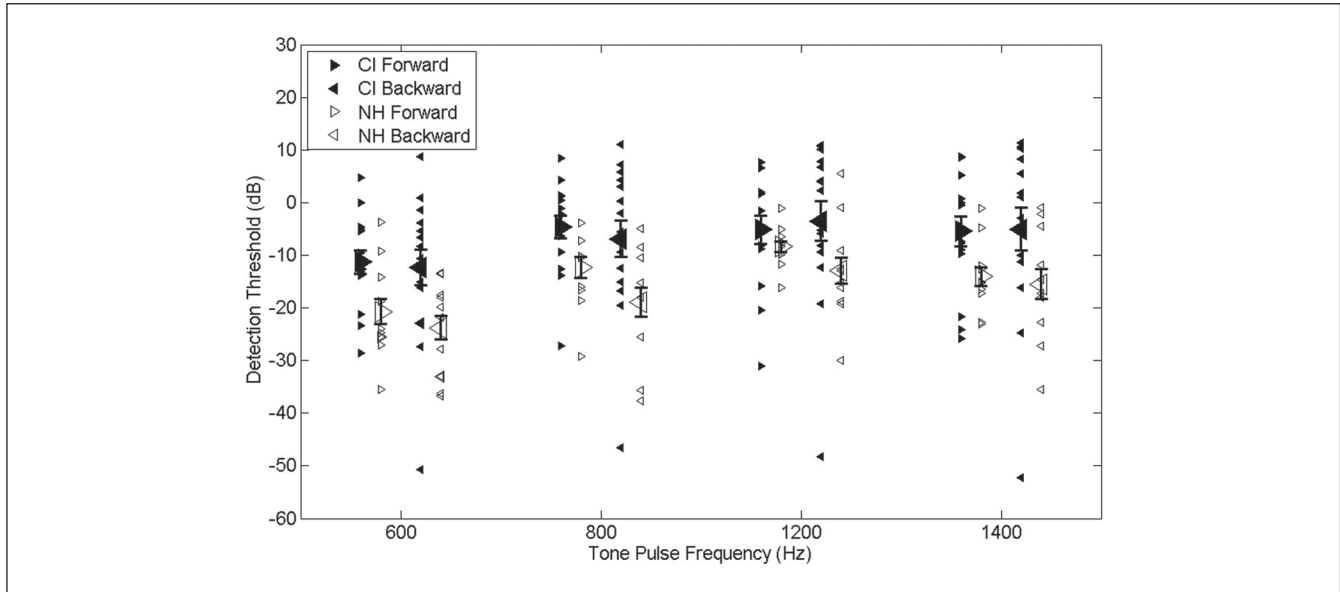


Figure 2. Forward- and backward-masked thresholds as a function of frequency. Detection thresholds (relative to 1000-Hz tonal masker level) are plotted for four different tone-pulse frequencies. Larger symbols and error bars represent mean values and standard error of the mean, respectively.

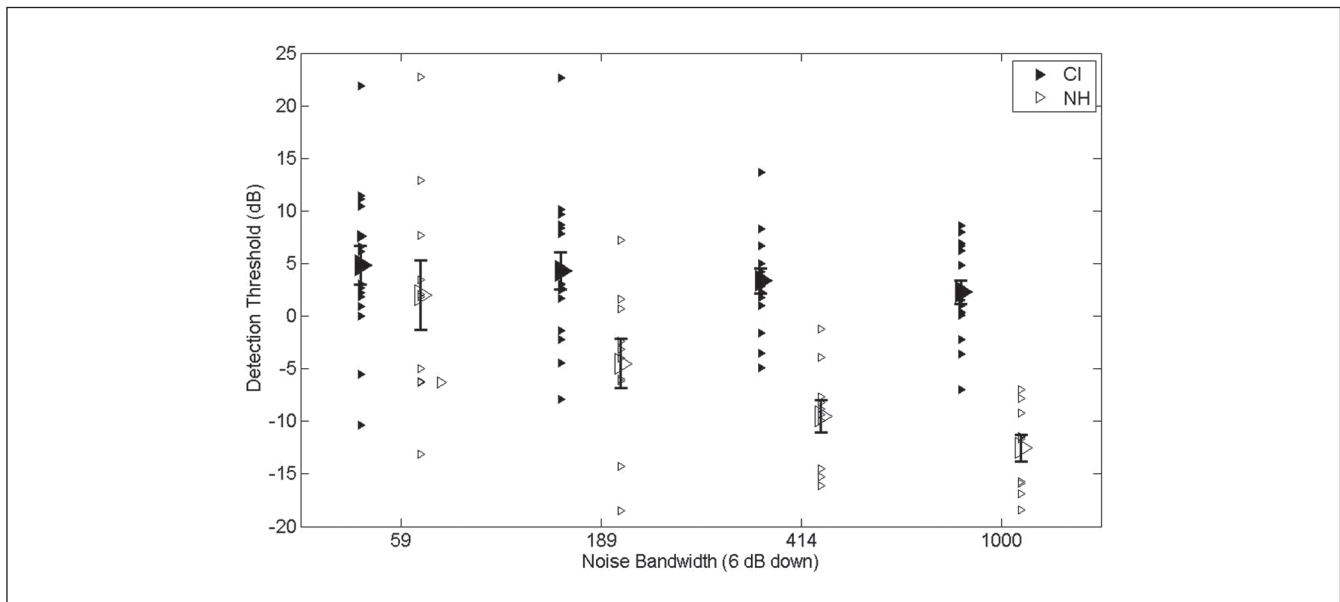


Figure 3. Detection thresholds plotted as signal to noise ratios in dB for 1000-Hz tones masked by narrow-band noise as a function of noise bandwidth. Larger symbols and error bars represent mean values and standard error of the mean, respectively.

noise bandwidth increases, but the improvement is greater for NH listeners. Average discrimination improves from 2.0 to -12.5 dB for NH listeners when increasing the noise bandwidth from 5.9% to 100% of the center frequency, while discrimination for CI listeners increases only from 4.4 to 2.0 dB.

Analysis of variance was computed on the tone-on-tone masking thresholds using group and condition as factors. Group was found significant, $F(1, 84) = 37.8, p < .001$, with

average thresholds for NH and CI listeners of -31.1 and -15.0 dB, respectively. Condition was weakly significant, $F(3, 84) = 2.1, p = .11$, and the interaction between group and condition was moderately significant, $F(3, 84) = 2.8, p = .047$. These effects can be observed in the plot of tone-on-tone masking thresholds versus tone masker frequency given in Figure 4. Listeners with normal hearing exhibit relatively lower detection limens for the higher frequency maskers compared to CI listeners.

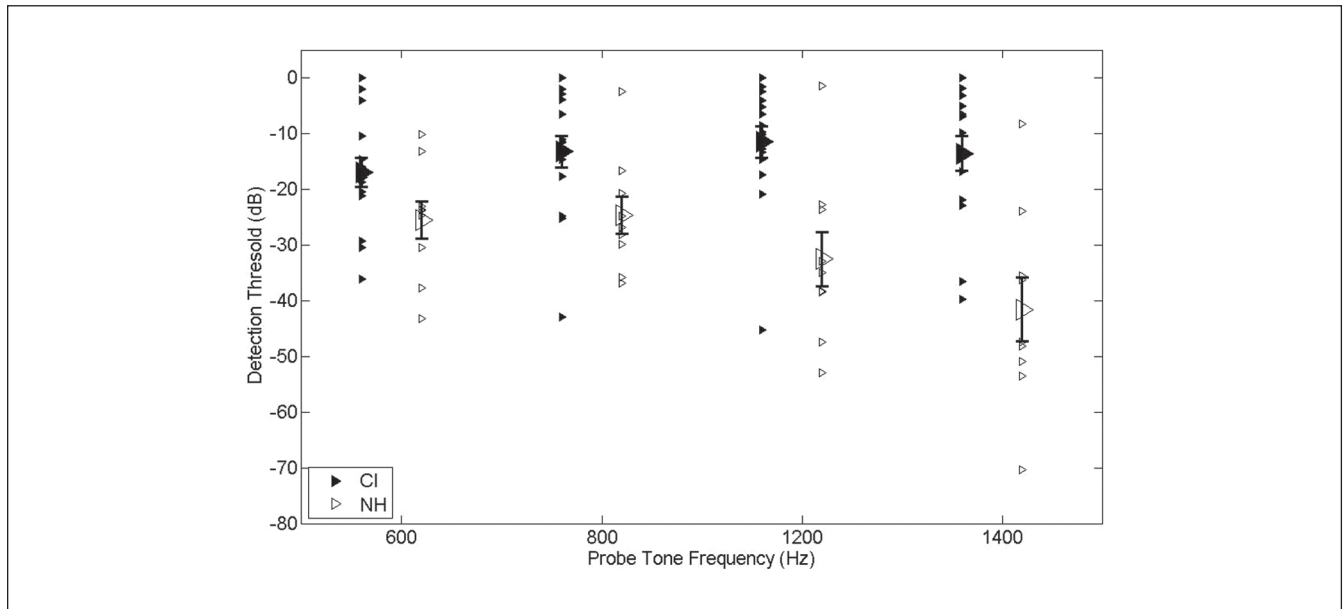


Figure 4. Masked detection threshold of probe tones in the presence of a 1000-Hz tonal masker as a function of the frequency of the probe. Larger symbols and error bars represent mean values and standard error of the mean, respectively.

Analysis of variance was computed on formant-on-formant masking thresholds using group and condition as factors. Group was significant, $F(1, 48) = 8.0, p = .007$, with an average NH threshold of -19.4 dB and an average CI threshold of -13.9 dB. Condition was not significant, $F(1, 48) = 0.8, p = .37$, but the interaction between group and condition was significant, $F(1, 48) = 11.0, p = .002$. The interaction effect is manifest in that NH listeners performed relatively better when the target was the lower frequency (-23.4 dB at 500 Hz) compared to the higher-frequency (-15.3 dB at 2000 Hz) formant. That effect was reversed in CI listeners who performed better at detecting the higher frequency (-11.6 dB at 2000 Hz) compared to the lower-frequency (-16.3 dB at 500 Hz) formant.

Many of the comparisons between measures within the pitch perception group were found significant for the CI listeners but not for the NH listeners. Among these were the correlation between forward- and backward-masking thresholds (Figure 5, Panel A) and the correlation between tone-in-noise masking and tone-on-tone masking (Figure 5, Panel B). For NH listeners, tone-on-tone masking thresholds were significantly correlated with the nonroved measures of tone-pip and harmonic complex-frequency discrimination. But in general, correlations among psychoacoustic measures were weak.

Phoneme Identification

The measures of phoneme perception are given in Appendix 4 and Figure 6. For vowel and consonant identification in quiet the average NH scores were 89.2% and 94.1% correct, whereas CI scores were 66.0% and 68.6% correct, respectively.

The best-performing CI listeners had comparable scores to the NH listeners with two of the implantees scoring within the NH range for vowels and five scoring within the NH range for consonants. Scores of the CI listeners, however, had a much wider range with a few listeners scoring less than 50% correct. This difference between groups was found significant in an analysis of variance computed separately on the vowel, $F(1, 27) = 15.9, p < .001$, and consonant, $F(1, 24) = 18.9, p < .001$, scores using group as a main factor (all ANOVAs for speech in quiet used scores converted from percent-correct to rationalized arcsine units: Studebaker, 1985).

For identification in quiet, consonant and vowel scores (Figure 6a) were significantly correlated ($r = .84$) for the CI listeners but much less ($r = .11$) for the NH listeners. This result is not surprising since CI listeners have a much wider range of performance (for vowels: 27.5%-88.1% correct, for consonants: 24.5%-87.1% correct). So listeners who performed poorly on the vowel measure generally performed poorly on the consonant measure; however, for the NH listeners, performance had a smaller range (for vowels: 86.3%-91.9% correct; for consonants: 80.0%-97.9% correct), so any inherent performance difference was not apparent.

Speech reception thresholds (SRTs) for gated noise are plotted as a function of the SRTs in stationary noise in Figure 6b. In general, identification of vowels and consonants deteriorated more rapidly in noise for CI listeners. For vowels, the average SRT in stationary noise for NH listeners was -13.9 dB compared to -7.6 dB for CI listeners. In gated noise, average SRT improved to -38.6 dB and -13.1 dB for NH and CI listeners, respectively. The masking release was thus 24.7 dB for NH listeners and 5.6 dB for CI listeners.

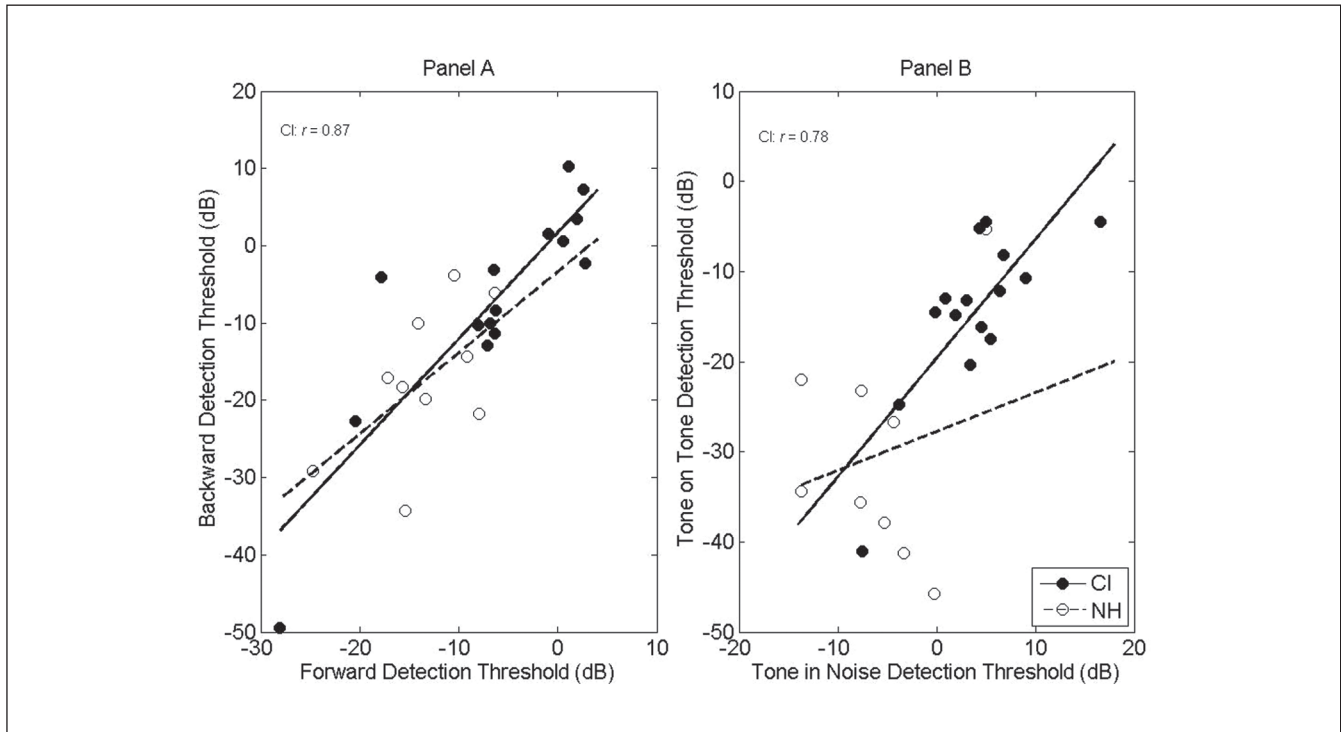


Figure 5. Forward-masked thresholds for tones as a function of backward-masked thresholds for tones (Panel A) in the presence of a 1000-Hz tonal masker. Thresholds are the target to masker ratios in dB. Detection thresholds for tones masked by 1000-Hz tones as a function of detection thresholds for 1000-Hz tones masked by narrow-band noise (Panel B). Thresholds are the signal to noise ratios.

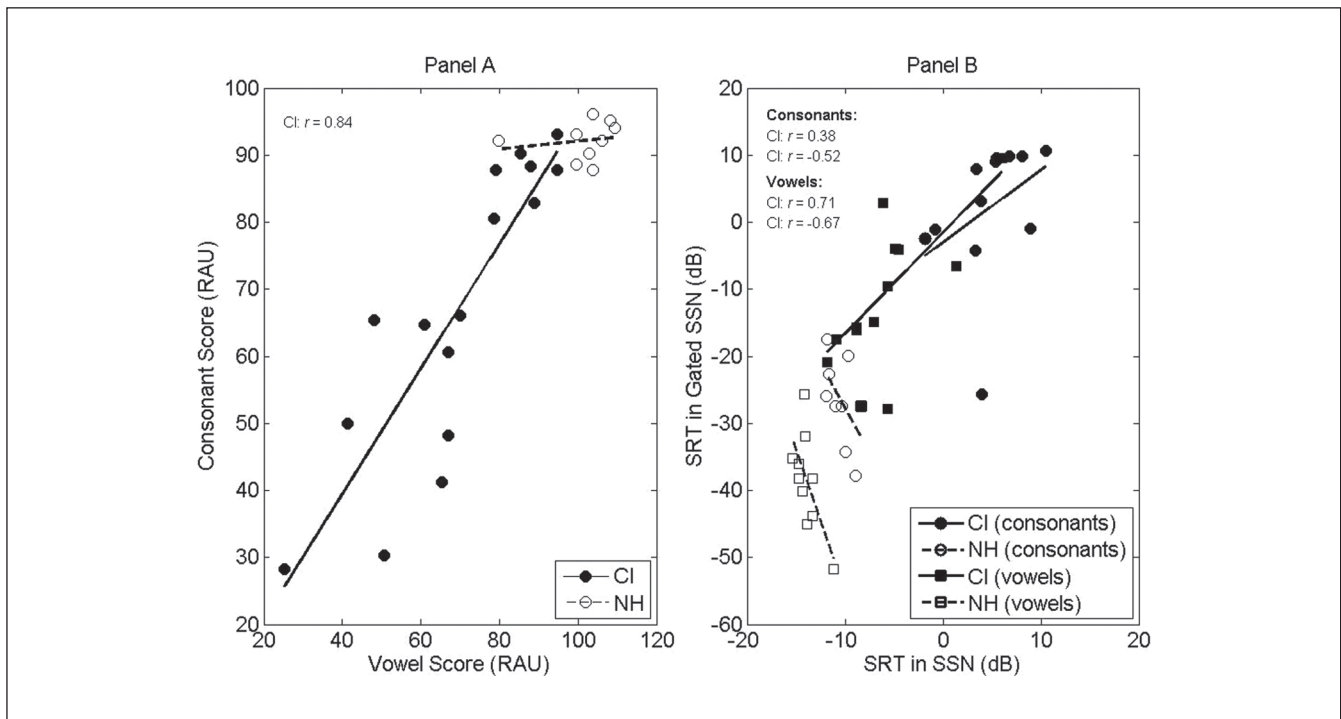


Figure 6. Vowel identification scores as a function of consonant identification scores (Panel A). Scores have been transformed to raus (Studebaker, 1985). The speech reception threshold (SRT) of vowels and consonants in steady-state noise as a function of SRTs in gated noise (Panel B).

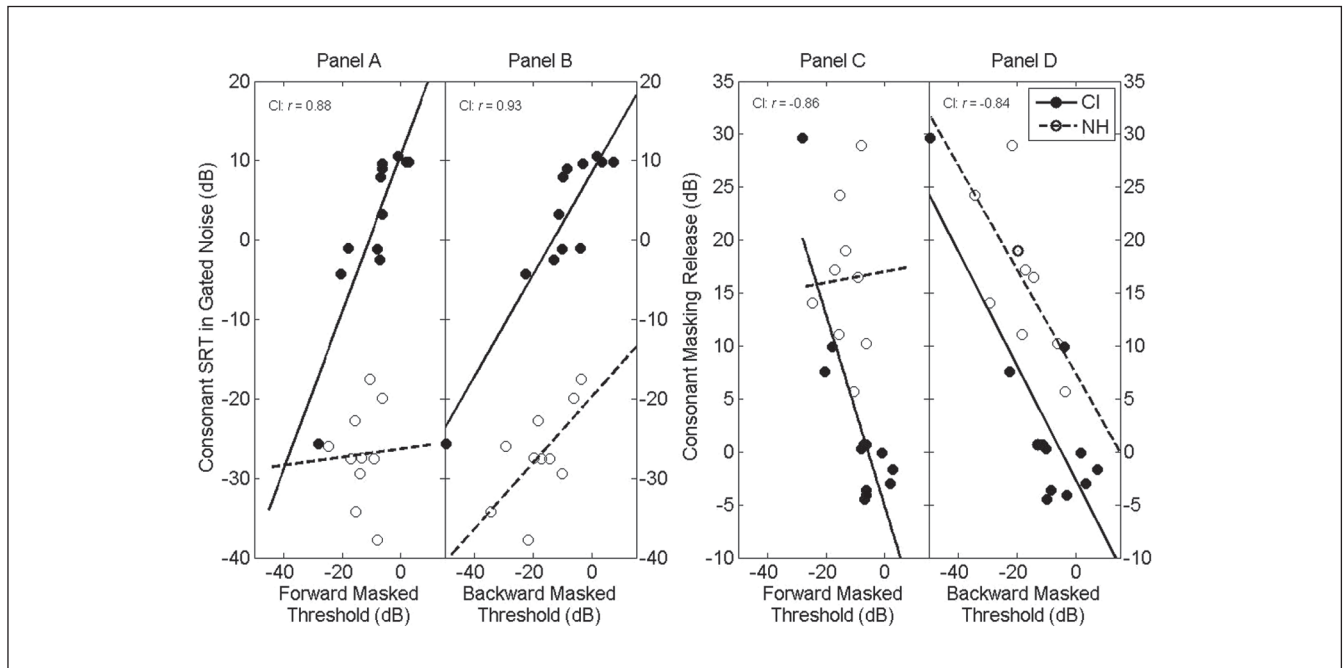


Figure 7. Consonant SRT in gated noise as a function of forward- (Panel A) and backward- (Panel B) masked thresholds. Consonant masking release as a function of forward- (Panel C) and backward- (Panel D) masked thresholds.

Analysis of variance computed on the vowel identification in noise measures using group and noise type (i.e., stationary or gated) as factors confirmed that group, $F(1, 50) = 93.6$, $p < .001$; noise type, $F(1, 50) = 57.5$, $p < .001$; and the interaction, $F(1, 50) = 26.1$, $p < .001$, were each significant.

Similar results were found for consonant identification in noise. The average SRT in stationary noise for NH listeners was -10.4 dB compared to 4.3 dB for CI listeners. In gated noise, average SRTs improved to -27.0 and 1.1 dB for NH and CI listeners, respectively. The masking release was thus 16.6 dB and 3.2 dB for NH and CI listeners, respectively. Analysis of variance computed on the consonant identification in noise measures using group and noise type (i.e., stationary or gated) as factors confirmed that group, $F(1, 50) = 170$, $p < .001$; noise type, $F(1, 50) = 26.3$, $p < .001$; and the interaction, $F(1, 50) = 17.3$, $p < .001$, were each significant.

Psychoacoustic and Speech Reception Measures

There were no significant correlations between psychoacoustic and phoneme identification measures for NH listeners. There were, however, significant correlations between psychoacoustic and phoneme identification measures for CI listeners. Both forward- and backward-masking thresholds were significantly correlated to consonant identification in gated noise ($r = .88$ and $.93$, respectively), plotted in Figures 7a and 7b. That is an increase in either forward or backward-masking threshold was associated with an increase in the

difficulty of perceiving speech in noise. Furthermore, both forward- and backward-masking thresholds were significantly correlated to masking release for consonants ($r = -.86$ and $-.84$, respectively), plotted in Figures 7c and 7d. Forward- and backward-masking thresholds were also significantly correlated to masking release for vowels ($r = -.79$ and $-.85$, respectively).

Tone-on-tone masking measures were also correlated to masking release for consonants ($r = -.92$) and moderately so for vowels ($r = -.77$; Figure 8). Formant-on-formant masking was correlated with vowel ($r = -.76$) and consonant ($r = -.79$) identification scores in noise (Figure 9). That is an increase in formant-on-formant masking threshold was associated with an increase in the difficulty of perceiving speech in quiet.

Finally, it is interesting to note that although we found that there were no other significant correlations between psychoacoustic and speech reception measures based on our significance criterion, all of the pitch perception measures had moderately high correlations (between $.59$ and $.76$) with consonant perception in speech-shaped noise. This may be indicative of the importance of place of articulation cues for consonant identification; however, the evidence is simply not clear since the pitch perception data were highly variable as a result of using untrained listeners.

Listener Age and Duration of Implant Use

Analysis of variance was computed on individual measures of CI performance after dividing listeners based on age and

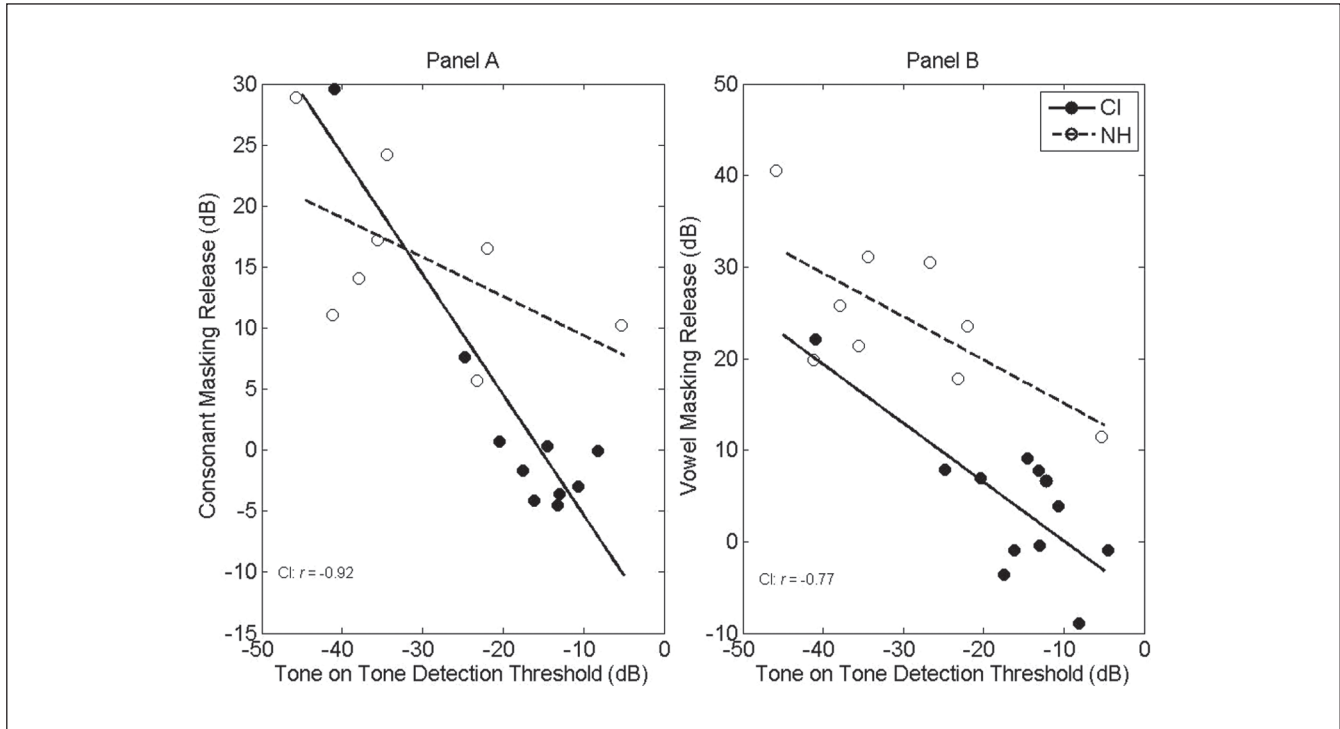


Figure 8. Masking release for consonants (Panel A) and vowels (Panel B) as a function of the detection threshold for tones masked by a 1000-Hz tone.

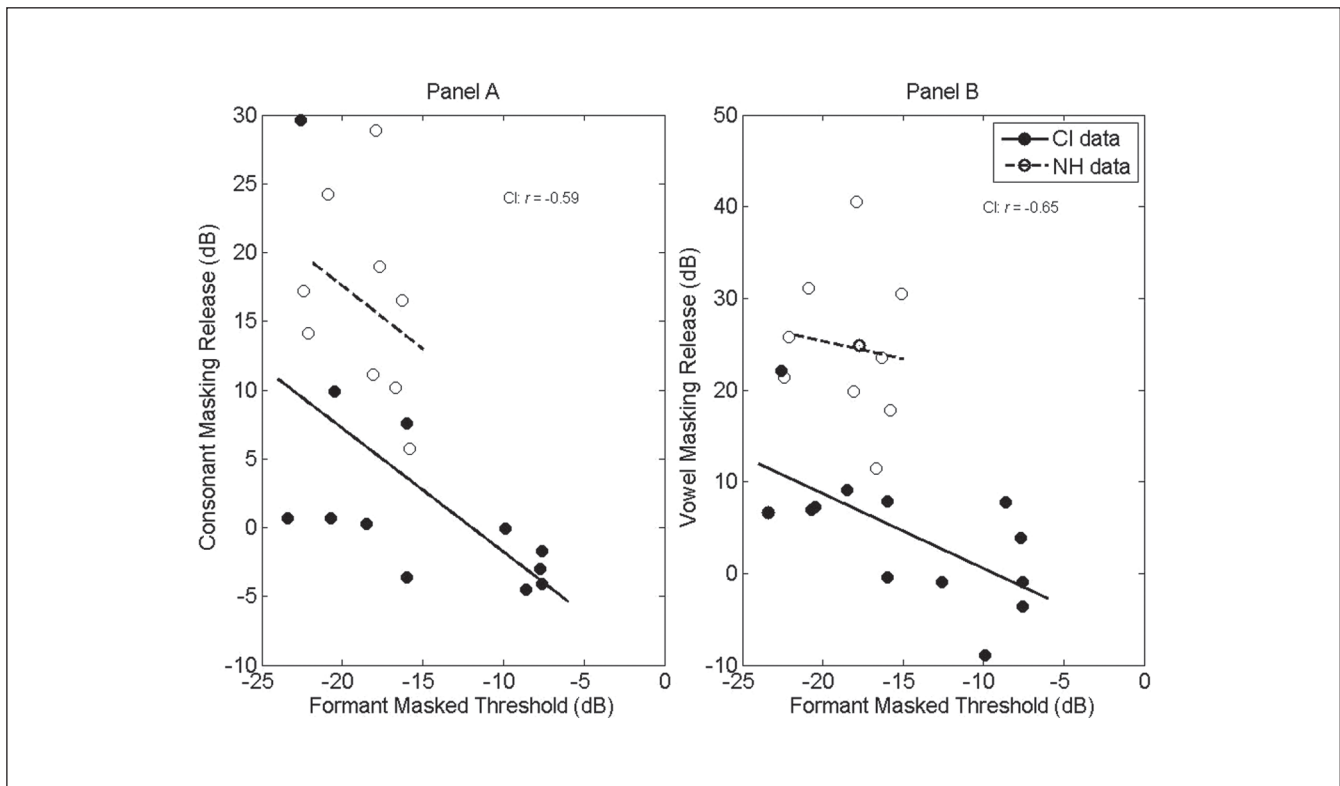


Figure 9. Masking release for consonants (Panel A) and vowels (Panel B) as a function of detection thresholds for formants masked by a formant.

duration of implant use. Dividing listeners as being older or younger than 55 years of age and computing analysis of variance on each measure yielded only one significant effect of group. It was found that the group of CI listeners younger than 55 performed better on the pitch perception tasks ($p = .002$) than the older listeners. For duration of use, all listeners had at least 1 year of implant experience. Dividing listeners as having more or less than 3 years of experience and computing analysis of variance on each measure yielded no significant effects of implant use.

Discussion

The purpose of this study was to examine the utility of using psychoacoustic measures as correlates of phoneme identification measures. The study design used protocols such that each measure could be obtained relatively quickly (in 5 to 10 min). The motivation for doing this is that it would be useful to make measurements quickly during clinical visits. However, doing so ignores the role of auditory practice on psychoacoustic tasks, so the results need to be considered with that in mind.

A primary goal of our study was to relate performance on psychoacoustic tasks that probe aspects of acuity such as pitch discrimination, intensity discrimination, and spectral-temporal masking to measures of vowel and consonant identification in CI and NH listeners. Although a number of significant correlations between psychoacoustic and speech tasks were observed for the CI listeners, no such instances were present in the NH data. The fact that correlations were observed with CI listeners and not for NH listeners may reflect the fact that CI listeners exhibit a wider range of scores on the psychoacoustic and speech reception measures than the NH listeners. That is, all of these measures will have a degree of variability in results that can be attributed to factors such as attention and simple guessing; the range of responses for CI listeners is wide enough that correlations can be observed despite such variability in results not related to audition. Whereas for the NH listeners, variations in results related to audition are relatively small (especially observed for the speech reception measures), so strong correlations are not observed.

For the CI listeners, measures of speech reception were not well correlated either with measures of pitch perception (with the exception of a significant correlation between tone-pip frequency DLs and vowel reception in gated noise) or intensity discrimination. One possible explanation for the weak correlation of speech with pitch perception abilities may lie in the fact that our relatively rapid assessments did not allow for asymptotic performance to be achieved on the pitch perception tasks. It is possible that any relationship that does exist may require a more detailed approach to making pitch measurements that allow a true psychoacoustic limit

to be reached before correlations will emerge with speech reception.

Similarly, psychoacoustic measures of intensity perception were found not to be correlated with speech reception measures. For the roved and nonroved pure-tone intensity discrimination, there may have been procedural learning aspects (similar to those in the pitch perception tasks) that limited listener performance and that might obscure a possible relationship. Given the recent report of strong correlation between a temporal-modulation detection task and speech reception in CI listeners (Fu, 2002), we were surprised that there was no correlation found between our measures of modulation detection and our measures of speech reception. The methods used by Fu required a more time-consuming series of measures at multiple intensity levels across listeners' dynamic ranges and additionally employed CI implants with similar signal-processing strategies. The lack of correlation using our abbreviated method may be a procedural learning effect, may be a consequence of using fewer modulation frequencies and stimulus levels as conditions, or may be related to having a variety of CI implant signal-processing strategies.

In contrast to the lack of correlations between measures of pitch and loudness discrimination with speech reception measures, we found the spectral-temporal masking measures to have moderate to strong correlation with phoneme identification. Gap detection tended to be the weakest predictor of speech reception performance, but even for it correlation was as high as .63 for consonant reception in quiet and in speech-shaped noise. Vowel and consonant identification in quiet were well predicted by formant-on-formant masking thresholds and SRTs for consonants in gated noise were correlated with forward and backward-masking thresholds. Several psychoacoustic measures were predictive of masking release in CI listeners for both consonants and vowels, with these measures including forward masking, backward masking, and tone-on-tone masking.

The NH listeners outperformed the CI listeners on all psychoacoustic and phoneme identification tasks except for intensity discrimination measures and gap detection. What is interesting, though, is that although significant correlations were observed between psychoacoustic and speech reception measures for the CI listeners, no such correlations were observed for the NH listeners. Of primary note is the large performance difference when comparing vowel and consonant identification in noise for CI versus NH listeners. Cochlear implantees had average SRTs of -7.6 dB on vowel identification and 4.3 dB on consonants identification for an 11.9 -dB difference. Part of this difference probably reflects the fact that the consonant task may be more difficult in that it has more alternatives from which to identify. But arguably more of the difficulty of the task arises from purely acoustic features in that consonants are relatively shorter in duration

and softer in acoustic intensity than vowels. It is interesting to note how CI listeners performed on vowel and consonant identification in noise relative to NH listeners. The corresponding SRTs in noise for NH listeners were -13.9 and -10.6 dB for vowels and consonants, respectively, for a relatively smaller difference of 3.3 dB. Furthermore, it should be noted that two of the CI listeners had measured SRTs less than -12 dB, which is only 2 dB worse than the observed average value for the NH listeners. However, the best-observed SRT for CI listeners on consonants was -4.1 dB, which is 6.5 dB worse than the NH average.

Of the psychoacoustic measures, forward and backward masking had the highest correlations with the observed measures of phoneme identification in CI listeners only. Coarsely, there are physical similarities in these two psychoacoustic measures with consonant identification. Consonant perception requires hearing of relatively brief, generally softer in intensity, components either preceding or following a sustained component. This analogy provides a rationale for the observed high correlations. The weakest of the correlations between these measures was in considering vowel identification in static noise with which neither forward ($r = .12$) nor backward masking ($r = .15$) was well correlated. But the remaining comparisons of speech measures with forward- and backward-masking measures were relatively high compared to the other psychoacoustic measures as summarized in Table 2.

It has been documented that NH listeners can understand speech better in the presence of fluctuating, as opposed to stationary, noise maskers (Festen & Plomp, 1983; Peters et al., 1998). In contrast, it has generally been found that CI listeners receive little or no benefit from fluctuating rather than stationary noise maskers (Cullington & Zeng, 2008; Donaldson & Nelson, 2000; Fu & Nogaki, 2004, 2005; Ihlefeld et al., 2010; Qin & Oxenham, 2003; Nelson & Jin, 2004; Nelson, Jin, Carney, & Nelson, 2003; Nelson, Nittrouer, & Norton, 1995; Nelson, Van Tasell, Schroder, Soli, & Levine, 1995; Stickney, Zeng, Litovsky, & Assmann, 2004). These prior studies primarily focused on sentence reception; in the present study, the focus was on phoneme identification. We found a large benefit for NH listeners. For vowels, NH listeners had an average SRT of -13.9 dB in stationary noise and -38.6 dB in gated noise for an average benefit of 24.7 dB. For consonants, NH listeners had an average SRT of -10.4 dB in stationary noise and -27.0 dB in gated noise for an average benefit of 16.6 dB. For both vowel and consonant materials the CI listeners had, on average, a small masking release. For vowels, CI listeners had an average SRT of -7.6 dB in stationary noise and of -13.1 dB in gated noise for an average masking release of 5.5 dB. For consonants, CI listeners had

an average SRT of 4.3 dB in stationary noise and of 1.1 dB in gated noise for an average masking release of 3.2 dB. While this masking release is relatively small compared to NH listeners, there is a small average benefit, with a high degree of variability, among CI listeners. Studies using noise-excited vocoding to simulate CI perception have found that masking-release increases with the number of channels in the vocoder simulation. The argument has been made (e.g., Fu & Nogaki, 2005) that increasing the effective number of spectral channels in CI listeners may lead to increases in masking release, and more generally, to improved speech reception in background noise (Fu et al., 1998; Munson & Nelson, 2005).

Li and Loizou (2008, 2009) provide a summary of factors affecting masking release in fluctuating noise. They argued that the first step in identifying phoneme patterns is to perceive acoustic events that demarcate the boundaries between vowels, consonants, and glide segments. They demonstrated that a clear presentation of obstruent portions of speech was sufficient to produce a release of masking in fluctuating noise for CI vocoded speech. This finding led them to argue that envelope compression and reduced dynamic range in CI sound processing smears the acoustic features to an extent that makes landmark identification untenable.

Our study yielded a psychoacoustic framework for the above argument. In particular, listeners with the best performance on measures of spectral-temporal masking tasks exhibited the greatest release of masking. Arguably, enhanced spectral-temporal resolution contributes to an improved ability to identify acoustic events (Chen & Loizou, 2010) and then to further characterize the phoneme structures accurately.

We emphasize that the methodology of our study was chosen to determine individual measures that could be conducted in less than ten minutes. The decision to make measurements with minimal training is relevant to any clinical test. Even with limited training, each condition required 20 to 30 min of testing. The downside of this approach is that it does not provide listeners with sufficient training to reach optimal performance on any of the tasks. Despite the lack of training on the measurements obtained here, correlations were nonetheless observed between certain psychoacoustic and speech tasks in the CI listeners. In particular, our results suggest that psychoacoustic measures of backward and forward masking may have predictive power for masking release for vowels and consonants. Future research on improved processing strategies for cochlear implants may include, for example, relatively quick measurements of forward masking to assess these strategies.

Appendix 1. Frequency Difference Limens (DLs). Normal-Hearing (NH) Data Averages and Ranges Are Shown (All Averages Were Calculated in a Logarithmic Scale).

	CI1R	CI2R	CI3L	CI3R	CI5R	CI6L	CI7R	CI9R	CI10L	CI10R	CI11L	CI11R	CI12L	CI13L	CI15L	CI18L	NH average	NH range
Condition	Pure-tone frequency DLs (%)																	
500 Hz	1.5	2.2	11.0	9.2	14.6	12.7	9.1	7.7	2.1	4.1	10.1	11.8	32.5	66.0	25.6	0.5	4.7	0.4-31.0
1000 Hz	0.5	0.6	8.5	13.8	7.3	2.5	13.4	76.8	2.9	2.3	21.1	27.0	17.0	49.1	34.1	0.3	5.1	0.3- 41.3
2000 Hz	1.7	1.4	1.7	8.7	5.3	1.3	29.1	6.3	2.6	6.6	11.3	88.8	18.9	30.4	5.1	0.2	3.7	0.3-18.8
	Pure-tone frequency DLs roved (%)																	
500 Hz	5.2	2.5	36.9	12.2	23.6	11.6	24.4	20.0	5.0	6.2	7.4	23.6	33.9	54.7	26.9	4.2	8.7	0.9-35.0
1000 Hz	4.7	6.6	9.7	16.0	31.9	14.3	22.0	74.9	3.3	7.5	13.9	72.6	71.2	79.7	61.3	5.4	11.2	1.2-39.5
2000 Hz	8.3	0.4	17.2	9.0	46.9	7.5	18.0	10.5	7.7	4.5	16.0	88.8	53.2	69.3	19.2	7.8	6.7	1.2- 37.9
	Tone-pip frequency DLs (%)																	
500 Hz	2.0	1.9	40.8	6.7	22.8	7.3	14.2	29.0	4.9	3.3	26.9	22.0	30.5	88.8	11.6	3.8	2.3	0.3-21.9
1000 Hz	1.6	4.1	13.0	12.3	13.9	1.9	48.5	15.7	4.5	5.0	55.7	6.3	45.6	43.6	43.0	0.5	2.0	0.2-14.7
2000 Hz	3.7	1.3	41.3	9.8	19.1	9.3	21.1	21.1	12.2	4.6	61.3	67.7	8.4	85.6	18.9	1.4	1.9	0.3-17.2
	Tone-pip frequency DLs roved (%)																	
500 Hz	5.3	2.3	46.9	15.0	22.4	11.0	32.4	27.8	7.3	11.6	15.5	26.4	21.5	29.8	8.6	6.2	5.9	0.8-38.2
1000 Hz	11.1	3.0	8.0	13.6	22.6	9.7	26.3	16.9	6.3	10.8	32.3	31.6	44.0	72.3	46.6	3.0	8.0	1.7-29.8
2000 Hz	18.2	3.0	26.4	19.1	35.7	7.8	18.9	18.1	3.9	10.6	62.1	73.9	30.0	38.5	20.2	5.3	6.9	0.8-49.9
	Harmonic complex frequency DLs (%)																	
110 Hz	3.5	0.8	20.2	23.3	8.4	12.2	41.1	15.1	6.6	11.3	15.2	22.5	52.2	45.6	40.8	2.1	3.0	0.2-45.6
220 Hz	1.9	3.1	19.8	7.9	11.8	5.6	21.8	22.5	16.0	29.9	13.2	83.1	44.7	45.2	48.3	2.1	4.3	0.6-33.6
440 Hz	8.3	62.0	20.2	4.5	13.2	21.0	9.9	23.0	0.9	27.2	3.9	62.7	35.0	57.6	46.4	16.4	3.6	0.6-23.3
	Harmonic complex frequency DLs roved (%)																	
110 Hz	2.3	27.5	34.1	12.6	32.4	25.6	32.8	9.7	9.7	13.2	29.8	78.3	77.6	25.3	3.0	2.3	7.2	1.3-48.7
220 Hz	7.4	25.3	15.9	47.9	27.8	24.2	51.2	28.4	29.6	29.9	86.8	37.6	77.3	47.5	5.4	7.4	7.8	1.2-43.6
440 Hz	48.1	30.8	32.8	84.9	29.8	42.9	46.4	74.0	48.9	11.1	67.8	60.8	78.6	37.8	49.6	48.1	5.9	1.4-43.4

Note: CI = cochlear implant.

Appendix 2. Intensity Difference Limens (DLs). Normal-Hearing (NH) Data Averages and Ranges Are Shown (All Averages Were Calculated in a Logarithmic Scale).

	CI1R	CI2R	CI3L	CI3R	CI5R	CI6L	CI7R	CI9R	CI10L	CI10R	CI11L	CI11R	CI12L	CI13L	CI15L	CI18L	NH Average	NH range
Condition	Pure-tone intensity DLs (dB)																	
0 dB	2.91	0.95	2.08	1.31	7.36	4.21	2.07	4.98	3.39	3.35	2.92	3.46	2.52	6.24	4.02	0.93	2.95	1.06-5.84
-6 dB	2.74	1.55	1.45	3.91	3.78	1.81	2.30	2.13	4.30	2.34	1.99	2.35	4.10	6.68	2.55	1.23	2.10	1.04-5.48
-12 dB	4.64	1.58	2.60	2.91	4.98	6.23	1.21	4.14	2.83	6.78	1.74	1.19	7.33	9.16	2.49	4.64	2.34	1.03-4.64
	Pure-tone intensity DLs roved (dB)																	
0 dB	4.64	4.48	6.24	1.39	5.24	3.46	4.19	4.85	9.11	8.54	4.19	7.95	5.64	4.64	4.72	4.39	4.46	2.46-6.33
-6 dB	7.23	3.14	8.71	2.42	3.95	1.98	1.83	5.00	8.35	10.68	4.61	7.68	5.74	7.23	7.85	6.68	4.28	1.23-7.23
-12 dB	4.64	1.93	5.49	8.63	5.05	3.60	2.81	3.31	5.48	8.36	1.41	5.11	8.98	4.64	2.03	7.13	4.00	2.57-6.23
	Modulation detection (%)																	
10 Hz	24.69	15.08	38.38	89.01	16.69	~	6.38	14.72	38.56	68.10	18.16	24.69	25.54	24.69	38.22	14.37	17.41	10.92-35.62
100 Hz	36.95	70.45	36.22	34.43	70.82	~	5.76	40.61	82.43	26.01	6.90	36.95	70.65	36.95	60.42	11.93	14.80	3.15-36.95

Note: CI = cochlear implant.

Appendix 3. Detection Thresholds for Spectral-Temporal Masking. Normal-Hearing (NH) Data Averages and Ranges Are Shown (Averages for gap Detection Were Calculated in a Logarithmic Scale).

	CI1R	CI2R	CI3L	CI3R	CI5R	CI6L	CI7R	CI9R	CI10L	CI10R	CI11L	CI11R	CI12L	CI13L	CI15L	CI18L	NH average	NH range
Condition	Gap detection (µs)																	
	6.56	3.93	10.98	9.53	12.17	9.27	6.76	9.17	6.68	6.84	6.10	16.55	14.75	7.24	7.83	8.09	6.17	5.05-7.93
	Forward masking (dB)																	
600 Hz	-12.6	-13.1	-9.57	4.67	-13.9	-21.2	-13.8	-5.08	-23.3	-11.8	-11.0	~	-0.10	-4.66	-5.43	-28.6	-20.45	-35.5 to 3.64
800 Hz	1.32	-6.32	-1.16	4.23	8.39	-12.6	-9.40	-2.22	-13.8	-4.52	-4.88	~	1.15	-2.49	0.51	-27.2	-12.18	-29.2 to 3.88
1200 Hz	-8.15	-4.97	6.57	1.93	7.66	-15.8	-4.80	-8.75	-20.5	-1.54	-4.58	~	6.65	-1.63	1.72	-31.0	-8.25	-16.1 to 1.16
1400 Hz	-8.95	-7.71	8.63	-0.51	8.77	-21.7	0.78	-9.63	-24.2	-7.42	-4.58	~	-0.10	5.12	5.26	-25.8	-13.86	-23.1 to 1.04
	Backward masking (dB)																	
600 Hz	-22.9	-10.6	8.77	0.84	-8.25	-16.1	-12.5	-3.90	-27.3	-15.0	-11.9	-13.1	-1.47	-6.56	-5.35	-50.7	-23.59	-36.7 to 13.4
800 Hz	-7.26	-6.89	11.01	7.12	-2.05	-15.1	-16.7	0.23	-19.5	-12.5	-9.39	-5.56	5.77	3.06	4.35	-46.6	-18.75	-37.7 to 4.87
1200 Hz	-5.32	-12.3	10.16	10.95	4.06	6.79	-5.85	-4.28	-19.3	-8.12	-9.42	10.72	3.98	7.86	2.28	-48.3	-12.83	-29.9 to 5.53
1400 Hz	-16.2	-11.3	11.3	10.3	-2.90	8.25	-5.11	-4.28	-24.8	-9.95	-2.90	10.51	5.52	1.82	1.07	-52.3	-15.42	-35.5 to 0.92
	Tone in noise (dB)																	
1/12 octave	7.62	0.90	1.87	2.70	21.86	-0.02	6.62	3.08	-5.50	6.15	2.22	11.11	11.43	7.59	10.48	-10.4	2.01	-13.1 to 22.7
1/4 octave	10.11	1.72	2.70	4.77	22.66	7.85	-1.40	2.99	-4.44	8.67	-2.20	2.46	9.68	8.36	3.08	-7.92	-4.50	-18.5 to 7.26
1/2 octave	6.69	-3.51	1.78	8.33	13.64	4.23	2.16	4.97	-1.63	1.01	3.49	3.85	8.33	2.82	2.32	-4.92	-9.51	-16.1 to 1.2
1 octave	1.01	0.30	1.51	6.26	8.02	0.36	4.88	6.91	-3.57	-2.20	0.11	2.78	6.68	8.62	2.02	-7.00	-12.54	-18.4 to 6.96
	Tone on tone (dB)																	
600 Hz	-17.9	-20.5	-18.7	-17.5	-10.4	~	-14.9	-21.2	-29.3	-30.5	-16.0	-1.99	-14.7	-16.2	-4.02	-36.1	-25.51	-43.2 to 10.1
800 Hz	-11.5	-11.6	-6.47	-24.8	-3.94	~	-6.00	-11.4	-12.7	-14.6	-25.3	-2.93	-17.7	-11.1	-2.08	-43.0	-24.71	-36.9 to 2.43
1200 Hz	-9.70	-12.7	-17.4	-14.7	-1.66	~	-27.6	-10.2	-20.9	-13.4	-4.05	-6.55	-5.21	-2.43	-8.62	-45.3	-32.57	-53.0 to 1.49
1400 Hz	-9.82	-13.1	-16.9	-13.0	-1.93	~	-4.20	-21.9	-36.5	-22.9	-6.76	-6.47	-5.11	-3.16	-7.03	-39.7	-41.56	-70.3 to 8.21
	Formant on formant (dB)																	
500 Hz	-20.8	-15.5	-3.96	-8.15	-11.3	-17.4	-4.01	-0.90	-8.1	-17.9	-17.3	-17.6	-9.87	-16.8	-4.63	-20.7	-23.45	-30.9 to 18.6
2000 Hz	-26.0	-21.4	-3.90	-7.03	-13.9	-23.6	-13.2	-14.3	-23.9	-23.5	-14.8	-3.25	-5.53	-3.78	-22.2	-24.5	-15.33	-25.2 to 8.23

Note: CI = cochlear implant.

Appendix 4. Measures of Phoneme Identification. Normal-Hearing (NH) Data Averages and Ranges Are Shown (Averages for Consonant and Vowel Identification in Quiet Were Calculated in a RAU Scale).

CI1R	CI2R	CI3L	CI3R	CI5R	CI6L	CI7R	CI9R	CI10L	CI10R	CI11L	CI11R	CI12L	CI13L	CI15L	CI18L	NH Avg.	NH Range
Vowel score (% correct)																	
86.3	88.1	27.5	50.0	66.3	80.6	86.3	40.6	86.7	82.5	66.9	29.4	61.3	65.6	48.1	90.0	89.2	86.3 to 91.9
SRT for vowels in SSN (dB)																	
-10.9	-11.8	~	6.0	-4.8	-8.8	-7.1	-4.9	1.4	-8.8	-4.5	~	-5.7	-6.1	~	-5.7	-13.9	-15.5 to 11.2
SRT for vowels in gated SSN (dB)																	
-17.5	-20.9	~	9.6	-3.9	-16.1	-14.9	-4.0	-6.5	-15.7	-4.1	~	-9.6	2.8	~	-27.8	-38.6	-51.7 to 25.7
Vowel masking release (dB)																	
6.6	9.1	~	-3.6	-0.9	7.3	7.8	-0.9	7.9	6.9	-0.4	~	3.9	-8.9	~	22.1	24.7	11.5 to 40.5
Consonant score (% correct)																	
91.0	84.5	24.6	40.8	47.9	78.8	79.2	66.3	86.5	87.1	70.8	50.8	67.9	61.7	67.9	91.0	94.1	80.0 to 97.9
SRT for consonants in SSN (dB)																	
-1.84	-0.8	~	8.1	~	8.9	3.4	5.5	3.3	3.9	5.4	~	6.8	10.5	~	4.0	-10.4	-11.9 to 8.4
SRT for consonants in gated SSN (dB)																	
-2.52	-1.1	~	9.8	~	-1.0	7.9	9.6	-4.3	3.2	9.0	~	9.8	10.6	~	-25.6	-27.0	-37.8 to 17.5
Consonant masking release (dB)																	
0.68	0.3	~	-1.7	~	9.9	-4.5	-4.1	7.6	0.7	-3.6	~	-3.0	-0.1	~	29.6	16.6	5.7 to 28.9

Note: CI = cochlear implant; SSN = speech-shaped noise; SRT = speech reception threshold.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Source of Support: National Institutes of Health, NIDCD.

References

- Boothroyd, A., Mulhearn, B., Gong, J., & Ostroff, J. (1996). Effects of spectral smearing on phoneme and word recognition. *Journal of the Acoustical Society of America*, *100*, 1807-1818.
- Busby, P. A., & Clark, G. M. (1999). Gap detection by early-deafened cochlear implant subjects. *Journal of the Acoustical Society of America*, *105*, 1841-1852.
- Chen, F., & Loizou, P. C. (2010). Contribution of consonant landmarks to speech recognition in simulated acoustic-electric hearing. *Ear and Hearing*, *31*, 259-267.
- Cullington, H. E., & Zeng, F.-G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. *Journal of the Acoustical Society of America*, *123*, 450-461.
- Donaldson, G. S., & 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. *Journal of the Acoustical Society of America*, *107*, 1645-1658.
- Dreschler, W. A., & Plomp, R. (1980). Relation between psychophysical data and speech perception for hearing-impaired subjects. I. *Journal of the Acoustical Society of America*, *68*, 1608-1615.
- Dreschler, W. A., & Plomp, R. (1985). Relations between psychophysical data and speech perception for hearing-impaired subjects. II. *Journal of the Acoustical Society of America*, *78*, 1261-1270.
- Dubno, J. R., & Schaefer, A. B. (1992). Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners. *Journal of the Acoustical Society of America*, *91*, 2110-2121.
- Festen, J. M., & Plomp, R. (1983). Relations between auditory functions in impaired hearing. *Journal of the Acoustical Society of America*, *73*, 652-662.
- Fu, Q.-J. (2002). Temporal processing and speech recognition in cochlear implant users. *NeuroReport*, *13*, 1635-1639.
- Fu, Q.-J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology: JARO*, *6*(1), 19-27.
- Fu, Q.-J., Shannon, R. V., & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. *Journal of the Acoustical Society of America*, *104*, 3586-3596.
- Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *Journal of the Acoustical Society of America*, *113*, 2861-2873.
- Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear-implant listeners. *Journal of the Acoustical Society of America*, *118*, 1111-1121.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099-3111.
- Kidd, G. R., Watson, C. S., & Gygi, B. (2007). Individual differences in auditory abilities. *Journal of the Acoustical Society of America*, *122*, 418-435.
- Kishon-Rabin, L., Segal, O., & Algom, D. (2009). Associations and dissociations between psychoacoustic abilities and speech perception in adolescents with severe-to-profound hearing loss. *Journal of Speech, Language and Hearing Research*, *52*, 956-972.
- Li, N., & Loizou, P. C. (2008). The contribution of obstruent consonants and acoustic landmarks to speech recognition in noise. *Journal of the Acoustical Society of America*, *124*, 3947-3958.
- Li, N., & Loizou, P. C. (2009). Factors affecting masking release in cochlear implant vocoded speech. *Journal of the Acoustical Society of America*, *126*, 338-346.
- Litvak, L. M., Spahr, A. J., Saoji, A. A., & Fridman, G. Y. (2007). Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *Journal of the Acoustical Society of America*, *122*, 982-991.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*, 36-47.
- Munson, B., & Nelson, P. B. (2005). Phonetic identification in quiet and in noise by listeners with cochlear implants. *Journal of the Acoustical Society of America*, *118*, 2607.
- Nelson, P. B., & Jin, S.-H. (2004). Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, *115*, 2286-2294.
- Nelson, P. B., Jin, S.-H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, *113*, 961-968.
- Nelson, P. B., Nitttrouer, S., & Norton, S. J. (1995). "Say-stay" identification and psychoacoustic performance of hearing-impaired listeners. *Journal of the Acoustical Society of America*, *97*, 1830-1838.
- Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., & Levine, S. (1995). Electrode ranking of "place pitch" and speech recognition in electrical hearing. *Journal of the Acoustical Society of America*, *98*, 1987-1999.
- Parks, T. W., & Burrus, C. S. (1987). *Digital filter design*. New York, NY: John Wiley.

- Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear implant processing on speech reception in fluctuating maskers. *Journal of the Acoustical Society of America*, 114, 446.
- Shannon, R. V. (1989). Detection of gaps in sinusoids and pulse trains by patients with cochlear implants. *Journal of the Acoustical Society of America*, 85, 2587-2592.
- Shannon, R. V. (1990). Forward masking in patients with cochlear implants. *Journal of the Acoustical Society of America*, 88, 741-744.
- Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E., & Wang, X. (1999). Consonant recordings for speech testing. *Journal of the Acoustical Society of America*, 106, 71-74.
- Stickney, G. S., Zeng, F.-G., Litovsky, R., & Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *Journal of the Acoustical Society of America*, 116, 1081-1091.
- Strouse, A., Ashmead, D. H., Ohde, R. N., & Grantham, D. W. (1998). Temporal processing in the aging auditory system. *Journal of the Acoustical Society of America*, 104, 2385-2399.
- Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research*, 28, 455-462.
- Supin, A. Y., Popov, V. V., Milekhina, O. N., & Tarakanov, M. B. (1994). Frequency resolving power measured by rippled noise. *Hearing Research*, 78, 31-40.
- Surprenant, A. M., & Watson, C. S. (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *Journal of the Acoustical Society of America*, 110, 2085-2095.
- Thibodeau, L. M., & Van Tasell, D. J. (1987). Tone detection and synthetic speech discrimination in band-reject noise by hearing-impaired listeners. *Journal of the Acoustical Society of America*, 82, 864-873.
- van Rooij, J. C., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *Journal of the Acoustical Society of America*, 88, 2611-2624.
- van Rooij, J. C., Plomp, R., & Orlebeke, J. F. (1989). Auditive and cognitive factors in speech perception by elderly listeners. I: Development of test battery. *Journal of the Acoustical Society of America*, 86, 1294-1309.
- van Schijndel, N. H., Houtgast, T., & Festen, J. M. (2001a). The effect of intensity perturbations on speech intelligibility for normal hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 109, 2202-2210.
- van Schijndel, N. H., Houtgast, T., & Festen, J. M. (2001b). Effects of degradation of intensity, time, or frequency content on speech intelligibility for normal hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 110, 529-542.
- Watson, C.S. & Kidd, G.R. (2002). On the lack of association between basic auditory abilities, speech processing, and other cognitive skills. *Rehabilitation*, 23, 83-94.
- Won, J. H., Drennan, W. R., & Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *Journal of the Association for Research in Otolaryngology: JARO*, 8, 384-392.
- Zwolan, T. A., Collins, L. M., & Wakefield, G. H. (1997). Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. *Journal of the Acoustical Society of America*, 102, 3673-3685.