

Across-site patterns of modulation detection: Relation to speech recognition^{a)}

Soha N. Garadat^{b)}

Kresge Hearing Research Institute, Department of Otolaryngology, University of Michigan, Ann Arbor, Michigan 48109-5616

Teresa A. Zwolan

Hearing Rehabilitation Center, Department of Otolaryngology, University of Michigan, Ann Arbor, Michigan 48109-0312

Bryan E. Pfingst^{c)}

Kresge Hearing Research Institute, Department of Otolaryngology, University of Michigan, Ann Arbor, Michigan 48109-5616

(Received 6 July 2011; revised 17 February 2012; accepted 20 March 2012)

The aim of this study was to identify across-site patterns of modulation detection thresholds (MDTs) in subjects with cochlear implants and to determine if removal of sites with the poorest MDTs from speech processor programs would result in improved speech recognition. Five hundred millisecond trains of symmetric-biphasic pulses were modulated sinusoidally at 10 Hz and presented at a rate of 900 pps using monopolar stimulation. Subjects were asked to discriminate a modulated pulse train from an unmodulated pulse train for all electrodes in quiet and in the presence of an interleaved unmodulated masker presented on the adjacent site. Across-site patterns of masked MDTs were then used to construct two 10-channel MAPs such that one MAP consisted of sites with the best masked MDTs and the other MAP consisted of sites with the worst masked MDTs. Subjects' speech recognition skills were compared when they used these two different MAPs. Results showed that MDTs were variable across sites and were elevated in the presence of a masker by various amounts across sites. Better speech recognition was observed when the processor MAP consisted of sites with best masked MDTs, suggesting that temporal modulation sensitivity has important contributions to speech recognition with a cochlear implant. © 2012 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.3701879>]

PACS number(s): 43.66.Ts, 43.66.Fe, 43.71.Es [PNN]

Pages: 4030–4041

I. INTRODUCTION

In current multichannel cochlear implants (CIs), both spectral and temporal cues are encoded in the electrical signal, and both types of cues have been found to contribute to the level of speech recognition with a CI (Xu *et al.*, 2005). However, given that CI recipients are reported to have reduced access to spectral cues (Friesen *et al.*, 2001), temporal cues become more important (Kwon and Turner, 2001; Qin and Oxenham, 2003; Shannon *et al.*, 1995; Xu *et al.*, 2002; Xu *et al.*, 2005). In most CIs, current is delivered to the electrodes in the form of amplitude-modulated trains of biphasic pulses. The modulation waveforms are generated proportional to the temporal envelopes of bandpass-filtered components of the acoustic signal. Therefore, the extent to which CI recipients are able to efficiently process temporal-envelope information may relate to their overall success with cochlear implantation. In support of this hypothesis,

differences in temporal processing have been shown to be a key factor differentiating between good and poor implant users (Chatterjee and Shannon, 1998; Fu, 2002; Colletti and Shannon, 2005; Luo *et al.*, 2008). Given these findings, it is necessary to examine temporal-envelope acuity within listeners to gain insight into the individual variability in this population.

One commonly used measure of temporal-envelope processing is the modulation detection threshold (MDT), which, in the context of the current experiments, is a measure of sensitivity to temporal-envelope modulations of pulse trains as a function of modulation depth. Several studies have demonstrated a relation between speech recognition and temporal modulation sensitivity in CI users (e.g., Cazals *et al.*, 1994; Fu, 2002; Luo *et al.*, 2008). Studies further demonstrated a considerable variability in modulation sensitivity across listeners (e.g., Fu, 2002; Richardson *et al.*, 1998; Chatterjee, 2003) and also across stimulation sites within listeners (Pfingst *et al.*, 2008). These results suggest that modulation sensitivity is dependent, at least in part, on local conditions near the individual implanted electrodes. This is consistent with studies showing that pathology in the deaf cochlea varies along the cochlear length in a patient-specific manner (Johnsson *et al.*, 1981; Hinojosa and Marion, 1983; Khan *et al.*, 2005; Fayad *et al.*, 2009).

^{a)}A portion of this work was presented at the 2011 Meeting of the Association for Research in Otolaryngology, Baltimore, MD.

^{b)}Current address: Department of Hearing and Speech Sciences, University of Jordan, Amman, Jordan 11942.

^{c)}Author to whom correspondence should be addressed. Electronic mail: bpfingst@umich.edu

In multichannel implants, competing signals at nearby stimulation sites (channel interaction) may also serve to disrupt temporal acuity. Experiments using a modulation detection interference (MDI) paradigm (Richardson *et al.*, 1998; Chatterjee, 2003) have shed light on the extent to which temporal acuity is susceptible to interference from channel interaction. These experiments measured sensitivity to amplitude modulated stimuli in the presence of interleaved maskers presented on another electrode. The elevation in modulation sensitivity was attributed in these experiments to overlap in neural excitation patterns generated by interleaved stimulation of two sites. The width of these excitation patterns along the length of the cochlea probably also depends on conditions near the individual implant electrodes (Bierer and Faulkner, 2010). Therefore, it is possible that masked MDT measures reflect both temporal and spatial features of the stimulation sites. The extent of the spatial contribution to overall site acuity can perhaps be estimated by deriving the difference in MDTs with and without a masker (amount of masking). Overall, one can assume that measuring modulation sensitivity in the presence of a masker is more likely to be informative of across site differences in temporal acuity in a multichannel processor than would a simple measure of MDTs from single electrodes. If poor performing sites can be successfully identified, perhaps several strategies or approaches can be devised to overcome the detrimental effect on overall performance that could be introduced by these suboptimal sites.

To date, the use of therapeutic approaches based on treating stimulation sites with neural pathology and poor functional responses have received minimal attention. Although neural regeneration at those sites is possible (Shibata *et al.*, 2010), it has not yet been implemented clinically. An alternative approach is to avoid stimulation at the suboptimal sites and to emphasize stimulation at optimal sites. Although such an approach would result in fewer sites being stimulated, previous studies have demonstrated some improvement in CI users' performance once the aberrant sites were removed from speech processor programs (Zwolan *et al.*, 1997). In

general, this approach might have significant implications in the clinical domain where data across sites could be used to guide CI fitting by selecting the best stimulation sites.

In this study, MDT measures were used to examine temporal-envelope acuity in quiet and in the presence of an adjacent interleaved masker. The overall goals of these experiments were (1) to identify across-site patterns of MDTs obtained with single-site and two-site stimulation and (2) to examine the effect of mean temporal acuity on speech recognition within listeners by selecting sites with the best masked MDTs versus sites with the worst masked MDTs for the speech-processor MAP. It was hypothesized that modulation sensitivity would decrease in the presence of a masker and that the amount of this decrease would be site-dependent. It was further hypothesized that choosing stimulation sites with better masked MDTs would improve speech recognition performance by maximizing temporal acuity and minimizing the detrimental effects of channel interaction on temporal acuity.

II. METHODS

A. Listeners

Twelve postlingually deaf subjects (6 males and 6 females) participated in this study. One subject had a Nucleus CI25M implant, five has CI24R (CS or CA) implants, and six had CI24RE (CA) implants. Subjects ranged from 51 to 75 years of age. All subjects had at least 12 months of experience using their cochlear implant. Demographic information for each subject is provided in Table I. Subjects' scores on sentence measures using their clinical MAPs are provided as a baseline to indicate the subjects' relative abilities with their normal everyday processor MAPs. These data were obtained using procedures described in Sec. II D. All testing for the current experiment was completed in four 8-h sessions per subject with frequent breaks during testing. All subjects were compensated for their participation. The use of human subjects in this study was reviewed and approved by the University of Michigan Medical School Institutional Review Board.

TABLE I. Subject information.

Subject	Gender	Age (years)	Age at onsets of deafness (years)	Duration of deafness (years)	Duration of CI use (years)	HINT Scores (% correct)	CUNY RTS (dB SNR)	Implant type	Strategy	Etiology of deafness
S52	F	58	2	56	2	83	14.5	CI24R(CA)	ACE	Trauma
S60	M	70	60	10	7	98	2.0	CI24R(CS)	ACE	Hereditary
S67	M	69	59	10	10	92	6.5	CI24R(CS)	ACE	Hereditary
S69	M	69	60	9	5	100	1.2	CI24M	ACE	Noise exposure
S77	M	59	52	7	4.5	64	12.0	CI24R(CA)	ACE	Hereditary
S81	F	59	52	7	5	100	-1.5	CI24RE(CA)	ACE	Hereditary
S82	F	60	36	24	6	99	0.8	CI24RE(CA)	ACE	Unknown
S83	F	65	13	52	6	100	0.5	CI24R(CS)	ACE	Otosclerosis
S84	M	51	26	25	5	98	2.3	CI24RE(CA)	ACE	Hereditary
S85	F	63	30	33	4	99	2.0	CI24RE(CA)	ACE	Hereditary
S86	F	64	60	4	3	99	4.7	CI24RE(CA)	ACE	Hereditary
S87	M	75	72	3	2.5	96	3.8	CI24RE(CA)	ACE	Hereditary

B. Psychophysical procedures

1. Electrical stimulation hardware

To assure uniformity in the external hardware, a laboratory Freedom[®] speech processor (Cochlear Corp., Englewood, CO) was used in psychophysics and speech testing. Communication with the processor was controlled using custom software run on a PC. Sequences of frames were generated and sent to the processor via COCHLEAR NIC II (Nucleus Implant Communicator) software (Swanson, 2004). Testing was conducted inside a large, carpeted double-walled sound-attenuating chamber to maintain consistency in the testing environment across sessions and subjects. All stimuli were delivered directly to the subject's implanted receiver/stimulator by software control of the external processor; no acoustic stimulation was used for the psychophysical tests.

Monopolar stimulation (MP 1+2) was used for all measurements. The term "stimulation site" in this study is used to indicate where stimulation was delivered along the electrode array. For monopolar stimulation, the stimulation site corresponds to the number of the scala tympani electrode, where the most basal electrode is number 1 and the most apical electrode is number 22.

2. Measurements of T and C levels

Absolute detection thresholds (T levels) and maximum comfortable loudness levels (C levels) were obtained using COCHLEAR CUSTOM SOUND SUITE 3.1 software. The stimulus burst duration was 500 ms with phase duration of 50 μ s presented at a rate of 900 pps and 200 ms in between bursts. T levels were obtained using an adjustment method until the subjects were able to detect a soft sound. The level of stimulation was then increased to the loudest level judged by the subject as the most comfortable to listen to for long periods of time. All electrodes were loudness balanced using an adjacent-reference loudness balance design (Zwolan *et al.*, 1997). Dynamic range (DR) for each site was calculated as the difference between the T level and the C level.

3. Modulation detection thresholds

Stimuli were 500 ms trains of symmetric biphasic pulses with a mean phase duration of 50 μ s and an interphase gap of 8 μ s presented at a rate of 900 pps and at 50% of the DR using custom software. Pulse duration, rather than current amplitude, was modulated because finer control of stimulus charge can be achieved using pulse duration. The phase duration of the pulse was modulated by a 10 Hz sinusoid that started and ended at zero phase. The positive and negative phases of the pulses were modulated equally to maintain charge balance and the interphase gap was held constant.

A four-alternative forced choice paradigm was used to determine MDTs at 50% of the DR. On each trial, subjects were presented with four sequential observation intervals 500 ms in duration, separated by 500 ms intervals and marked by boxes on the computer screen. One of the four intervals was randomly chosen on each trial to contain the modulated signal while the remaining three intervals contained unmodulated pulse trains. Subjects were instructed to

choose the interval that sounded different from the other three intervals. Feedback was provided after each trial by marking the chosen interval as "C" for a correct response or "X" for an incorrect response. A two-down, one-up adaptive-tracking procedure (Levitt, 1971) was used, starting with a modulation depth of 50% and ending when 12 reversals were recorded. Modulation depth was increased or decreased in steps of 6 dB to the first two reversals, 2 dB to the next two reversals, and 1 dB for the next eight reversals. Each run was terminated after 12 reversals were obtained, and MDTs were defined as the average of the modulation depths at the last eight reversals.

Using this procedure, detection thresholds for depth of modulated pulse duration were determined for electrodes 2-21 in the subjects' electrode arrays in a random order. MDTs were obtained for single-site stimulation and in the presence of an interleaved unmodulated masker presented on the adjacent apical site (masked MDTs). Three MDTs were obtained for each site under the quiet and masked conditions using a different randomization each time; the average of these three values was used as the estimate of the MDT for each site. The average values were then used to calculate the amount of masking by subtracting MDTs in quiet from those in the presence of a masker.

C. Speech processor MAPs

Based on the masked-MDT data, the experimental speech processor was mapped using approaches that allowed comparisons of speech recognition when subjects used two MAPs that contrasted in across-site-mean modulation sensitivity. A total of 10 electrodes was used in each MAP where the first MAP contained electrodes with the best masked MDTs and the second MAP contained electrodes with the worst masked MDTs. Excluding the most basal and apical electrodes (electrodes 1 and 22), the remaining 20-electrode array was divided into five segments, each containing four stimulation sites. To maintain the full range of place-pitch representation, two sites were selected from each segment to form the 10-site experimental maps. For the best-MDT MAP, the two sites from each segment with the best masked MDTs were used. For the the worst-MDT MAP, the two sites with the worst masked MDTs were used. An example of site selection is illustrated in Fig. 1.

All subjects were fitted using the continuous interleaved sampling (CIS) processor strategy. In general, the CIS strategy uses interleaved stimulation and is designed to preserve temporal details in the speech signal envelope by using high-rate carrier pulse trains (Wilson *et al.*, 1991). Stimuli were symmetric biphasic pulses with a phase duration of 37 μ s and an interphase gap of 8 μ s. Standard clinical procedures were used to determine the T and C levels for each electrode as well as loudness balancing at the C level across electrodes. T and C levels for the processor fitting were determined using 500 ms pulse trains presented at a rate of 900 pps. To enable the CIS strategy in the 10-channel MAPs, the maxima number was set to 10. In general, the two MAPs were identical in all aspects except for selection of the stimulation sites. To ensure uniform performance across subjects in the sound

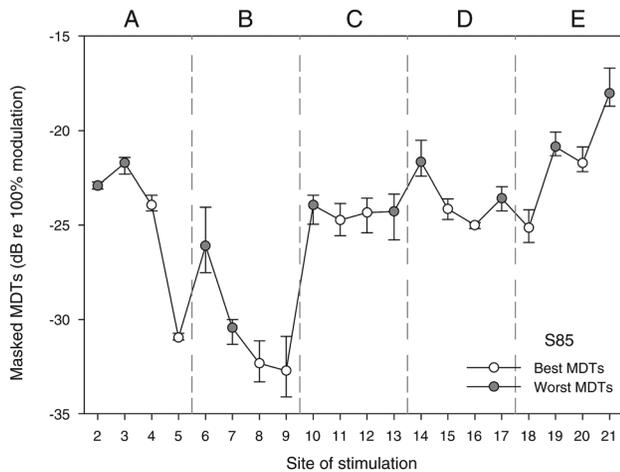


FIG. 1. Masked MDTs for stimulation sites 2-21 are displayed for one subject (S85) to demonstrate the site-selection approach used to construct the best-MDT MAP and the worst-MDT MAP. For each site, the mean MDT and the range for three repeated measures are shown. The electrode array (sites 2-21) was divided into 5 segments as shown (labeled A-E). From each segment, the two sites with the lowest (best) masked MDTs were selected to construct the best-MDT MAP (open symbols). In contrast, two electrodes with the highest (worst) masked MDTs were selected from each segment to construct the worst-MDT MAP (filled symbols).

field testing, a basic MAP was the standard fitting protocol in this study. Specifically, SmartSound options were deactivated; these included ADRO, beam, whisper, and auto sensitivity.

A sensitivity setting of 8 was used for all subjects; this setting was fixed and was disabled in the control unit to prevent the subjects from changing the setting during testing. However, subjects were instructed to adjust the volume setting on their processor to a comfortable level while listening to sentences from the HINT list presented in quiet and in noise; these sentences were not reused during speech testing. Several volume settings were tried with both MAPs until subjects found a setting that was comfortable for both MAPs, and the subjects were then instructed to maintain this setting throughout the testing session. Therefore, different volume settings were used across subjects but the volume setting used by each subject was the same for the two MAPs.

D. Speech measurements

Subjects' performances with the two MAPs were evaluated using four speech batteries. For each battery, testing for the two MAPs was completed in the same session and the order of testing was randomized across the two MAPs. Speech tests consisted of medial vowel and consonant discrimination as well as sentence recognition in quiet and in noise. HINT speech-shaped noise was used when speech recognition was evaluated in the presence of an acoustic masker. Speech testing was conducted in a carpeted double-walled sound booth (Acoustic Systems Model RE 242 S). Stimuli were passed through a Rane ME60 graphic equalizer and a Rolls RA235 35W power amplifier and presented through a loudspeaker positioned about 1 m away from the subject at 0° azimuth. All target stimuli were presented at 60 dB SPL with the exception of sentences in noise which

were adaptive. Speech levels were calibrated with a sound-level meter (Brüel & Kjær, Naerum, Denmark, type 2231). The sound-level meter microphone was positioned at 1 m away from the loudspeaker and 0° azimuth. A slow time setting and an "A" frequency weighting were used in the calibration.

Vowel stimuli (Hillenbrand *et al.*, 1995) consisted of 12 medial vowels presented to subjects with custom software in a /h/-vowel-/d/ context (head, hawed, head, who'd, hid, hood, hud, had, heard, hoed, hod, hayed). Using a 12-alternative forced-choice paradigm, subjects were instructed to click on the appropriate token on a computer screen to indicate their response selection. The 12 vowels were alphabetically arranged and presented in a grid on the computer screen. Two male talkers and two female talkers were selected from the recordings yielding a total of 48 stimuli per run. Stimuli were presented in quiet and in the presence of HINT speech-shaped noise using four different signal-to-noise ratios (SNRs) of +15, +10, +5, and 0. Tokens were presented in random order without replacement. Chance level on this test was 8.33% correct.

Consonant stimuli (Shannon *et al.*, 1999) consisted of 20 naturally spoken American English consonants presented in a consonant-/a/ context (ba, cha, da, fa, ga, ja, ka, la, ma, na, pa, ra, sa, sha, ta, tha, va, wa, ya, and za). Two talkers (1 male talker and 1 female talker) were used for a total of 40 presentations. Stimuli were presented in quiet and at SNRs of +15, +10, +5, and 0. A 20-alternative forced-choice paradigm was used, and subjects were instructed to click on one of the 20 choices on a computer screen to indicate their response. Chance level for the consonant test was 5% correct.

Recognition of words in sentences was measured in quiet using the Hearing in Noise Test (HINT) sentences (Nilsson *et al.*, 1994) and in HINT speech-shaped noise using City University of New York (CUNY) sentences. Subjects were instructed to verbally repeat their responses. The HINT sentences were obtained from the Minimum Speech Test Battery for Adult Cochlear Implant Users (House Ear Institute and Cochlear Corp., 1996). Lists of 10 sentences each were drawn at random without replacement from the 25 lists in the test corpus so that a different list was used for each repetition. A score was calculated based on the percentage of the correct words repeated per list.

For the CUNY sentences, an adaptive procedure was used to measure reception thresholds for sentences (RTS) in the presence of HINT speech-shaped noise. RTS is defined as the SNR needed to achieve 50% sentence intelligibility. Each sentence was scored by the experimenter as correct if the subject successfully repeated all the words in a given sentence or incorrect if the subject missed any of the words in the sentence. Based on the accuracy of the subject's response, the SNR was decreased (following a correct response) or increased (following an incorrect response) in 2 dB steps. The peak level of the combined signal (sentences plus noise) was held constant while the SNR changed from trial to trial. As noted in the preceding text, subjects were instructed to adjust their speech processor to a comfortable listening level and to not change the setting during testing.

All adaptive tracks were started at an easy SNR of 20 dB and were completed after 10 reversals were obtained. RTS was calculated as the average of the SNRs at the last 6 reversals.

All speech stimuli for the different speech measures were presented a single time only, and feedback was not provided. Each test was presented two times, and the results of the two different lists were averaged to yield speech recognition scores. Training was provided prior to the data collection for each measure. The training session for the consonant and vowel recognition consisted of a preview of the stimuli and then practice runs with feedback in quiet and in noise. For measures of sentence recognition, training was provided in the form of two runs in quiet and two runs in noise using a new list each time; feedback was not provided.

III. RESULTS

A. Across-site patterns of MDTs

Figure 2 shows MDT results for the 12 CI subjects with data from each subject shown in a separate panel. The open

symbols represent MDTs obtained in quiet, and the filled symbols represent the MDTs obtained in the presence of an unmodulated pulse train on the adjacent site. As evident from the figure, there was considerable variation in modulation sensitivity across sites of stimulation as well as across listeners. In addition, MDTs were elevated in the presence of the masker. Similarly, masked MDTs were variable across listeners and across sites of stimulation. To allow comparison of MDTs across listeners and in relation to other measures, the MDTs were averaged across stimulation sites for each listener; this is referred to as across-site mean (ASM). ASM values are reported in Fig. 2 (lower right corner of each panel) for the MDTs in quiet and for the masked MDTs. In addition, across-site variance (ASV) was calculated for each listener as reported in the following text.

In Fig. 3, the relations in ASMs (A) and ASVs (B) are plotted for the MDTs in quiet and in the presence of a masker. Results showed a strong significant correlation between MDTs in quiet and in the presence of a masker for ASMs ($r^2 = 0.97$, $P < 0.0001$) and for ASVs ($r^2 = 0.94$,

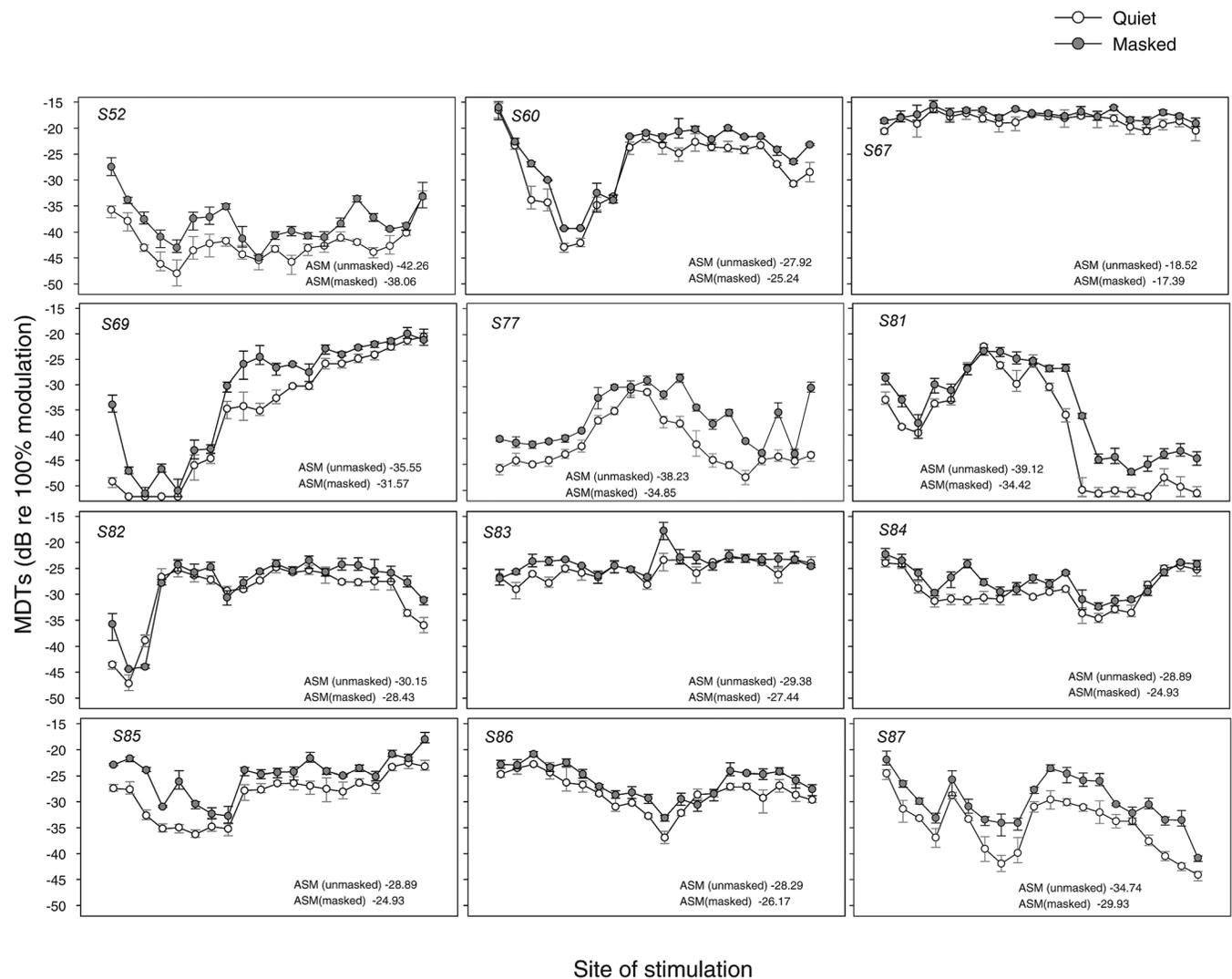


FIG. 2. MDT means and ranges are shown as a function of stimulation site. Each panel represents data from one subject with subject numbers shown in the upper left corner of each panel. Open symbols represent MDTs in quiet, and filled symbols represent MDTs in the presence of a masker on the adjacent apical electrode. Across-site means (ASM) are calculated separately for MDTs in quiet and masked MDTs and displayed for each subject in the lower right corners of the panels. Larger negative values indicate better performance.

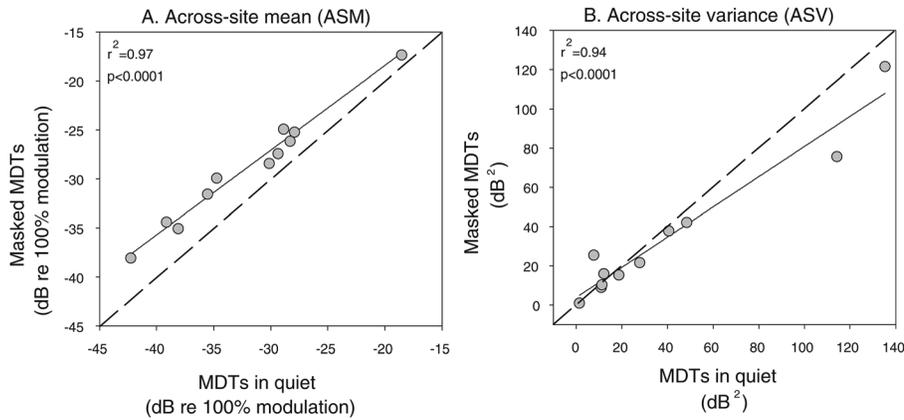


FIG. 3. Scatter plots are shown to illustrate the correlation in ASMs (left-hand panel) and ASVs (right-hand panel) between the MDTs in quiet (x axis) and masked MDTs (y axis) for each listener. The solid lines are best fit linear regression lines; r^2 and P values are shown in the upper left corners of the panels. In (A), all points are slightly above the dashed diagonal line, indicating that ASM masked MDTs were a little higher in all subjects than ASM MDTs in quiet.

$P < 0.0001$). These results demonstrated that subjects who exhibited good temporal acuity performed nearly as well in the presence of a masker as they did in quiet, suggesting that differences across listeners were primarily mediated by differences in temporal acuity.

To examine whether the variation in MDTs was systematic along the cochlear length for listeners, the electrode array was divided into five segments (see Fig. 1). Specifically, the five segments from most basal to most apical were A (sites 2-5), B (sites 6-9), C (sites 10-13), D (sites 14-17), and E (sites 18-21). For each segment, ASM and ASV of MDTs in quiet and masked MDTs were calculated as shown in Figs. 4(a) and 4(b). These data were separately subjected into one-way repeated measure ANOVA with segment as the within-subject factor. Results indicated no significant differences in ASM across segments, suggesting that across-listeners' differences in MDT patterns were related to subject-specific variation in conditions along the length of the electrode array and were not just due to normal physiological differences along the cochlear length. However, differences in ASV as a function of cochlear segment were significant for MDTs in quiet [$F(4, 11) = 3.142, P < 0.05$]

and in the masked conditions [$F(4, 11) = 3.370, P < 0.05$]. *Post hoc* analysis showed that for MDTs in quiet, ASV was significantly larger for segment A than for segment D ($P < 0.05$) and also significantly larger for segment B than for segment D ($P < 0.05$). For masked MDTs, ASV was significantly larger for segment A when compared to that for segments C and D ($P < 0.05$). These results, in general, suggest that there is greater variation in modulation sensitivity in the basal segment when compared to other segments along the tonotopic axis, which was perhaps related to variation in electrode position as discussed in the following text.

B. Across-site patterns of amount of masking

To quantify the elevation of thresholds in the presence of a masker, amount of masking was calculated for each site as (masked MDT – MDT in quiet). These results are shown in Fig. 5. A value of zero indicates no elevation in threshold when a masker was present, a positive value indicates that thresholds were elevated in the presence of a masker, and a negative value indicates that thresholds were lower in the presence of a masker. The amount of masking was variable

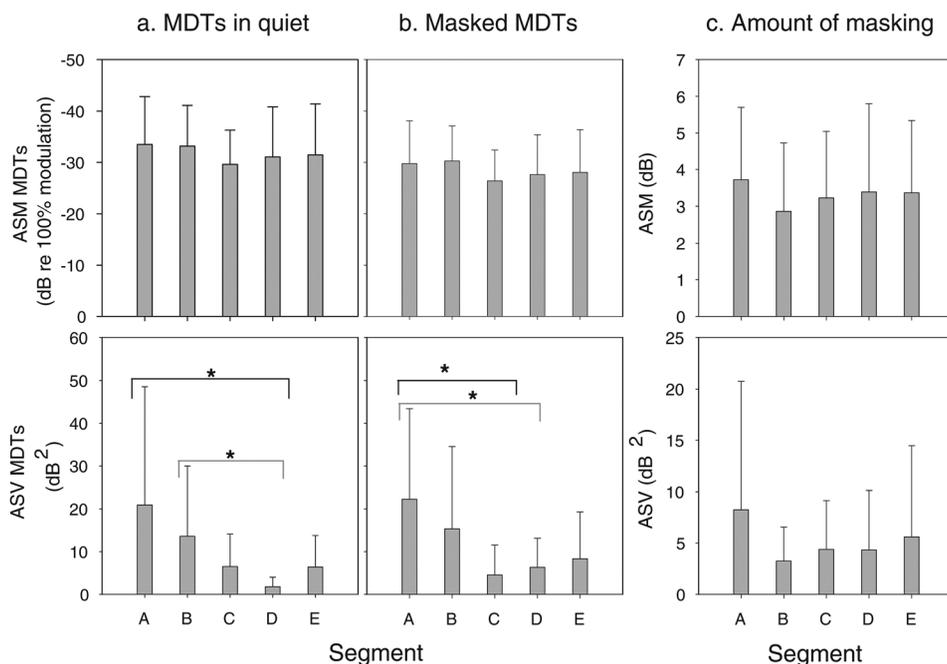


FIG. 4. ASMs and ASVs of MDTs are plotted as a function of place of stimulation. The electrode array was divided into 5 segments of four electrodes in each segment as shown in Fig. 1. ASM (upper panels) and ASV (bottom panels) are shown for MDTs in quiet (a), masked MDTs (b), and amount of masking (c). An asterisk represents a statistically significant difference ($P < 0.05$). The error bars represent standard deviations.

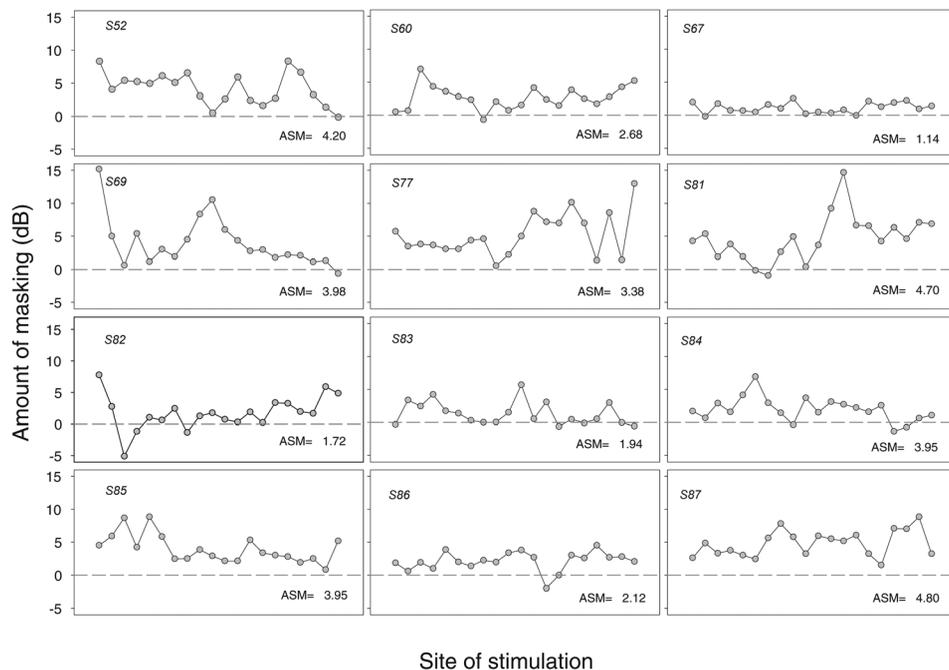


FIG. 5. The amount of masking in decibels (masked MDT minus MDT in quiet) is shown as a function of stimulation site. Each subject is represented in a different panel. The dashed line indicates no differences in MDTs when a masker was present; a positive value indicates an increase in MDTs in the presence of an interleaved masker on the adjacent site. ASM amount of masking is given for each subject at the lower right corner of each panel.

across stimulation sites, suggesting that these differences may be related to differences in the local conditions in the cochlea near the stimulation sites.

To examine whether differences in amount of masking were systematic along the cochlear length, the electrode array was divided into five segments, and ASM and ASV were calculated for each segment; results are shown in Fig. 4(c). ASM and ASV values for amount of masking were subjected to statistical analysis similar to that reported in Sec. III A. There were no significant differences in ASM or ASV across the five cochlear segments.

C. Comparisons of modulation sensitivity and amount of masking for the two MAPs

This study aimed to create two experimental MAPs with different mean temporal acuity in the presence of an unmodulated masker. Therefore, further analyses were conducted to examine the extent of differences in modulation sensitivity and in amount of masking between these two MAPs. For each subject, data were rearranged into two 10-site sets corresponding to the sites used to construct each MAP. ASM and ASV of MDTs in quiet and in the presence of masker as well

as amount of masking were calculated for each MAP; results are shown in Fig. 6. These data were compared using one-way repeated measures ANOVA with MAP assignment as the within-subject factor. Results of these analyses confirmed that the best-MDT MAP had overall better ASM modulation sensitivity in quiet [$F(1, 11) = 41.153, P < 0.0001$] and in the presence of a masker [$F(1, 11) = 103.522, P < 0.0001$] than that obtained with the worst-MDT MAP. Overall amount of masking for the best-MDT MAP was significantly less than that for the worst-MDT MAP [$F(1, 11) = 23.604, P < 0.01$]. On the other hand, the analyses showed that there were no differences ($P > 0.05$) between the two MAPs in the ASVs of MDTs in quiet or in the presence of a masker. These results suggest that although we reduced variability per segment, variability in modulation sensitivity remained large across the electrode array for each MAP. However, less variation across sites of stimulation (ASV) in amount of masking was obtained for the best-MDT MAPs when compared to that with the worst-MDT MAPs [$F(1, 11) = 5.230, P < 0.05$]. This is perhaps due to the fact that amount of masking was a derived value rather than an absolute value; therefore, variability would shadow differences between the two MAPs in amount of masking.

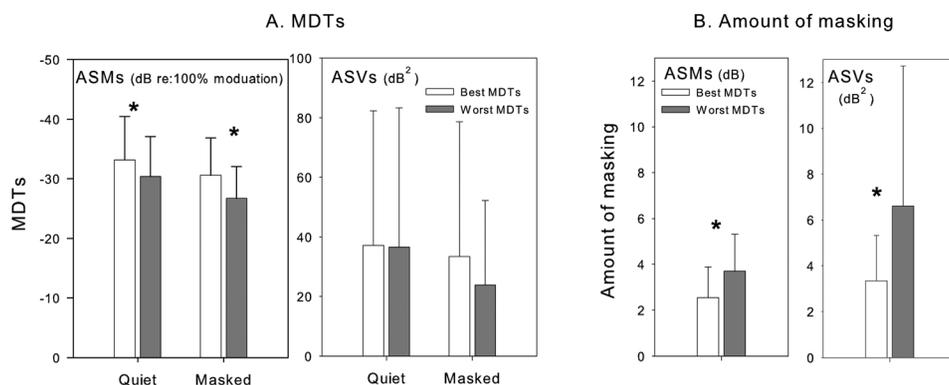


FIG. 6. ASM and ASV for MDTs (A) and amount of masking (B) are compared for the two sets of electrodes that were selected to construct the best-MDT MAP and the worst-MDT MAP. An asterisk represents a statistically significant difference between the two MAPs at P value of 0.05. The error bars represent standard deviations.

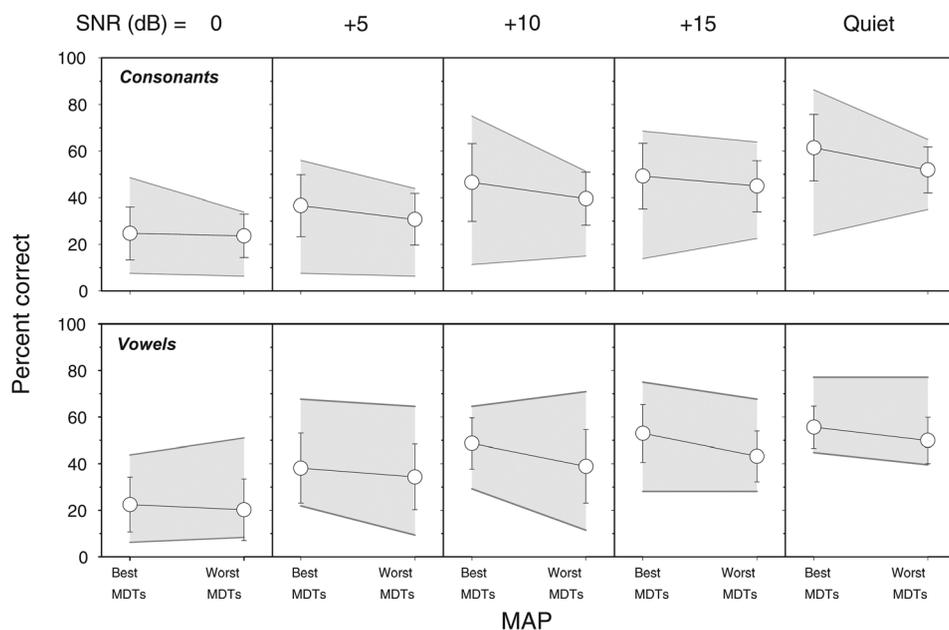


FIG. 7. Average discrimination percent correct scores for consonants (upper panels) and vowels (bottom panels) are shown as a function MAP. The shaded area represents the range of the scores across the 12 listeners. From left to right, the panels show discrimination scores as a function of increasing signal-to-noise ratio (SNR) in dB. The error bars represent standard deviations.

D. Comparisons of speech recognition for the two MAPs

To determine the effect of the site-selection procedure on speech recognition, the two MAPs were compared for the various speech recognition measures. Results from each of the different speech measures were subjected to a linear mixed model analysis with random subject effects. Several variables, such as gender, everyday mode of stimulation, and volume were different among the subjects and could potentially have confounded the effects examined in this study. For example, it was important to control for mode of stimulation (unilateral, bilateral) that subjects used in everyday listening to account for the concern when a bilateral user is forced to listen with one CI. In addition, given that it was important to conduct the speech tests at a comfortable listening level for each listener, this resulted in several volume control settings (6, 7, 8, and 9). It was felt necessary to account for this variable in the analysis given that temporal acuity measures improve with increased stimulation levels, which could perhaps obscure any potential differences in speech recognition between the two MAPs. Therefore, these variables were included as covariates in the analyses. None of the analyses showed that gender, everyday listening mode, or volume were predictors of performance on the speech measures. Furthermore, controlling for these factors did not affect the variance estimates, indicating that these three variables do not explain within-subject and between-subject variations. Given these results, analyses in this study were repeated without the inclusion of these variables and results from these analyses are reported in the following text.

The top row of Fig. 7 and Table II(a) present average scores for the 12 subjects for the consonant recognition test. Percent scores for the best-MDT MAPs and the worst-MDT MAPs compared in quiet and at SNRs of 0, +5, +10, and +15 are displayed in different panels. Data were subjected to a linear mixed model statistical analysis where the independent variables were MAP (best-MDT and worst-MDT) and SNR (0, +5, +10, +15, and quiet). Results indicated

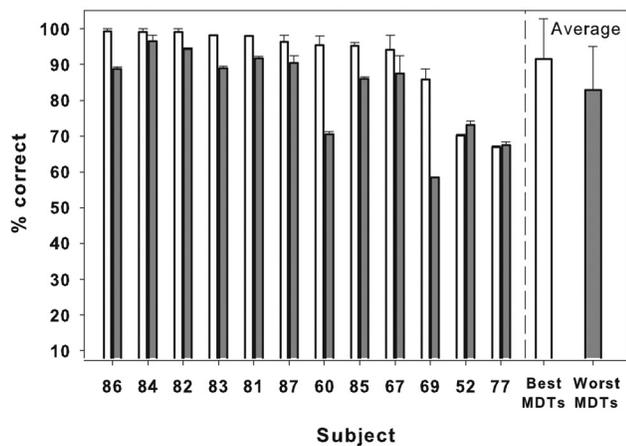
that consonant discrimination scores were better with the best-MDT MAP than that obtained with the worst-MDT MAP by an average of 5.6%; this was a significant improvement ($P < 0.0001$). Results further revealed that effect of SNR was also significant ($P < 0.0001$). Follow-up tests were conducted to evaluate mean differences in scores as a function of SNR; the Bonferroni procedure was used to control for Type I error. *Post hoc* pairwise *t*-tests showed that consonant discrimination scores linearly improved with increased SNR ($P < 0.0001$). However, there were no differences between scores for SNRs of +15 and +10 dB. The interaction between MAP and SNR was not significant.

A similar analysis was conducted for the vowel discrimination scores; results are shown in the bottom row of Fig. 7 and in Table II(b). Similar to the results for consonants, overall vowel discrimination was better with the best-MDT MAP than the worst-MDT MAP by 6% ($P < 0.0001$). Results further revealed that effect of SNR was also significant ($P < 0.0001$). *Post hoc* pairwise *t*-tests showed that scores of vowel discrimination linearly improved with increasing SNR ($P < 0.01$) up to +10 dB. There were no significant differences in scores obtained with SNRs of +15 and +10 dB and those obtained in quiet and SNR of +15 dB, but better scores were obtained in quiet than those obtained

TABLE II. Consonant and vowels average percent scores.

a. Consonants					
MAP	0 SNR	5 SNR	10 SNR	15 SNR	Quiet
Best MDT	33 ± 15	45 ± 17	52 ± 18	59 ± 17	70 ± 18
Worst MDT	29 ± 16	40 ± 14	50 ± 17	55 ± 17	64 ± 19
b. Vowels					
MAP	0 SNR	5 SNR	10 SNR	15 SNR	Quiet
Best MDT	39 ± 20	54 ± 19	62 ± 16	64 ± 17	69 ± 12
Worst MDT	45 ± 19	57 ± 16	64 ± 15	69 ± 15	72 ± 10

A. HINT in quiet



B. Adaptive CUNY in noise

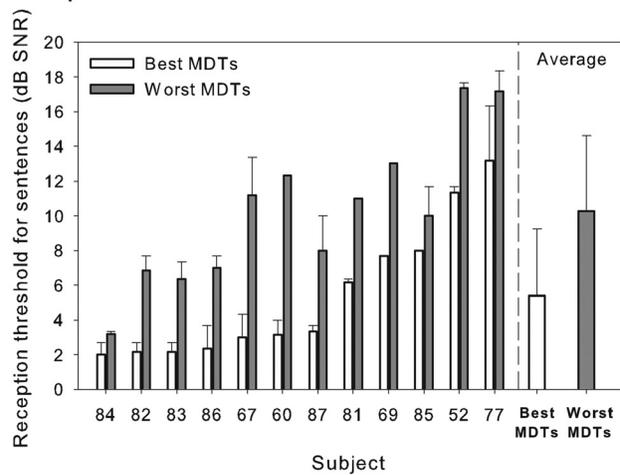


FIG. 8. Listeners' performance on sentence discrimination tasks is shown for HINT sentences in quiet (A) and for CUNY sentences in noise with adaptive SNRs (B). In (A), percent correct is measured and, therefore, the greater the score, the better the performance. In (B), reception thresholds for sentences are measured in decibel SNR, and hence, the smaller the number, the better the performance. Individual data as well as averages across all subjects are compared here for the best-MDT MAPs (light bars) and the worst-MDT MAPs (dark bars); error bars represent the range. Across-subject mean are shown in the right-hand portion of each panel; the error bars represent standard deviations.

with SNR of +10 dB ($P < 0.01$). The interaction between MAP and SNR was not significant.

It is also important to note that the residual and variance estimates of these analyses were significant ($P < 0.05$) for the consonant and vowel discrimination. This indicates that there were within- and between-listener variations in speech scores. These variations remained large even after controlling for the three variables indicated earlier, suggesting that these variables did not drive these within-subject and between-listener variations in the results. Rather, these variations were likely introduced by unexplained variables.

Sentence discrimination results are shown in Fig. 8(A) (HINT sentences in quiet) and in Fig. 8(B) (CUNY sentences in noise). For the HINT sentences in quiet, most subjects (except S52 and S77) showed higher percentages of correct responses for the best-MDT MAP when compared to responses obtained when using the worst-MDT MAP. For the CUNY sentences in noise, all subjects achieved 50% correct at lower (more difficult) SNRs with the best-MDT MAP than with the worst-MDT MAP. Thus, performance with the best-MDT MAP was better than that for the worst-MDT MAP for sentences in quiet and for sentences in noise for most of the subjects.

Speech scores from each test were separately subjected to a linear mixed model analysis with random subject effects; the independent variable was MAP. Results showed that subjects' performance on the best-MDT MAP was better than their performance on the worst-MDT MAP for both HINT ($P < 0.01$) and CUNY ($P < 0.0001$) tests. Specifically, the improvement in performance was 7% in the HINT sentences and 5 dB SNR in the CUNY sentences. In addition, results revealed that the residual estimate was significant for both HINT and CUNY measures. However, the between-subject component (subject variance) was significant for the CUNY measure but not for the HINT measure; this perhaps is due to a ceiling effect seen for the majority of subjects when HINT sentences were used.

IV. DISCUSSION

In this study, MDTs were surveyed from all electrodes to determine differences in temporal patterns of neural activation elicited at different sites in the cochlea. Differences across the electrode array were then used to construct new speech processor MAPs. Based on the assumption that speech recognition could be improved by eliminating suboptimal sites for stimulation, this study was conducted to determine the feasibility of a site-selection approach based on modulation sensitivity.

A. Across-site variation in modulation sensitivity

Consistent with previous reports, MDTs were highly variable across listeners and also across sites of stimulation within individual implant users (Fu, 2002; Colletti and Shannon, 2005; Pfungst *et al.*, 2008). These variations could be related to local irregularities in the vicinity of sites of stimulation resulting in poor electrode-neuron interfaces at some stimulation sites. It is also possible that these differences in MDTs could be induced, in part, by systematic differences along the tonotopic axis of the cochlea. The current study, however, does not directly support this latter hypothesis because MDTs did not change systematically as a function of length along the cochlear spiral (Fig. 4, top row) and because the patterns of MDTs along the cochlear spiral were different for each subject (Fig. 2). On the other hand, variation in modulation sensitivity across sites was found to be largest in the basal region (Fig. 4, bottom row). These results are perhaps related to more variable cochlear pathology in the basal region (Nadol, 1997) or to greater across-subject variation in intracochlear electrode array placement with respect to the modiolus in the basal region.

The present study further showed that MDTs were elevated in the presence of an unmodulated pulse train interleaved on an adjacent stimulation site within and across listeners. This is in line with previous studies demonstrating

similar elevation in MDTs in the presence of a masker (Richardson *et al.*, 1998; Chatterjee, 2003). Interestingly, overall patterns of masked MDTs were strongly correlated to patterns of MDTs in quiet as can be seen in Fig. 3. These results indicate that site to site variations in modulation sensitivity in quiet and in the presence of a masker were perhaps mediated by the same mechanisms underlying temporal acuity. One possible source is the alteration in the temporal response properties of neural activation that could be altered by local variations in pathology along the tonotopic axis. In support of this notion, it has been shown that the temporal features of the neurons can be altered by the loss of the myelin sheath (Tasaki, 1955) or a shift in the action potential generation site to a more distal point along the auditory nerve (Javel and Shepherd, 2000). The extent of this alteration could be stimulation-site specific depending on the local conditions.

Although patterns of MDTs were strongly correlated across the quiet and masked conditions in this study, as can be seen in Fig. 3, modulation sensitivity on some sites degraded to a greater extent than on other sites when a masker was present. The variations in the extent of this degradation were perhaps related to differences in the spatial properties across sites of stimulation that could be introduced by variables such as the anatomy of the cochlea and the geometry of the electrode array. Any pathological mechanisms that might be present would also alter the manner in which neurons interact, leading to increased channel interaction and reduced temporal resolution in some cases (Vollmer *et al.*, 1999; Vollmer *et al.*, 2000), which perhaps was reflected in this study by sites with larger amounts of masking, as shown in Fig. 5. In sum, the extent of the overlap in the neural excitation patterns at these sites may partially underlie the variability in masked MDTs.

B. Effect of modulation sensitivity on speech recognition

Overall, this current work is in agreement with previous studies suggesting that listeners' ability to encode and resolve temporal patterns of the electrical signal supports the dynamics of speech recognition (Eddington *et al.*, 1978; Wilson *et al.*, 1991; Shannon, 1992). Specifically, in this study, listeners achieved better consonant and vowel recognition scores with the best-MDT MAP (sites with the best masked MDTs) than those obtained with the worst-MDT MAP (sites with the worst masked MDTs). These results along with other work suggest that modulation sensitivity may serve as a predictive measure of subjects' performance on consonant and vowel recognition in multichannel CIs (Cazals *et al.*, 1994; Fu, 2002; Colletti and Shannon, 2005). Additionally, results showed that overall scores improved as a function of SNR as anticipated (e.g., Fu *et al.*, 1998; Friesen *et al.*, 2001). Although there was no statistically significant interaction found between MAP and SNR, in Fig. 7 one can notice that differences in percent scores between the two MAPs decrease as SNR decreases; this is most noticeable at 0 SNR. These results could be related to a floor effect at this adverse listening condition.

This study demonstrated that sentence recognition was also better with the best-MDT MAP compared to that obtained with the worst-MDT MAP. On average, the advantages of listening with the best-MDT MAP compared to the worst-MDT MAP were 7% for HINT sentences in quiet and 5 dB SNR for the CUNY sentences in noise. The advantages seen with the best-MDT MAP were more consistent across subjects for measures of sentence recognition than those found with the consonant and vowel measures. Additionally, differences in performance between the two MAPs were more consistent across subjects for the sentence recognition in noise test (CUNY) than for the sentences in quiet (HINT). These results perhaps suggest that temporal modulation sensitivity is more important in noise than in quiet. The present study further demonstrated that listeners did not necessarily have poor speech recognition with the worst-MDT MAP, but rather speech recognition was improved relative to the worst-MDT MAP when temporal acuity in each segment was optimized. Perhaps, these results support previous studies demonstrating the resilience of speech recognition under significantly degraded conditions (e.g., Van Tassel *et al.*, 1987; Shannon *et al.*, 1995). Previous studies have shown that CI users effectively use only four to eight channels (Friesen *et al.*, 2001); hence, the 10-channel MAPs used here should provide sufficient spectral cues. In the context of this argument, it is important to consider the possibility that differences in speech recognition between the two MAPs might be larger for conditions with fewer numbers of channels because temporal information is more salient in cases where spectral resolution is worse (e.g., Shannon *et al.*, 1995; Xu *et al.*, 2002; Xu *et al.*, 2005). Thus improving mean temporal acuity might be particularly useful for users with poor speech recognition because spectral resolution (i.e., the number of effective channels) for these users is greatly reduced (Friesen *et al.*, 2001). This remains an open question for future studies.

In summary, overall speech recognition performance was better when the best-MDT MAP was used relative to performance with the worst-MDT MAP. However, these results do not imply that the observed advantages for the best-MDT MAP would also apply when compared to the listeners' performance with their clinical MAP. A direct comparison of speech performance between these experimental MAPs and the clinical MAP is not that straightforward for a number of reasons including differences between the two MAPs in number of channels, speech-processing strategy, and the fitting parameters. These variables would create a major confound in interpreting the current data, so no such comparison was performed. However, sentence recognition results with the clinical MAP are provided in Table I as a baseline to show the relative abilities of these subjects to recognize speech with their normal everyday processor maps.

Consistent with previous reports, the current work further demonstrated a considerable variation in speech recognition scores across listeners (Gantz *et al.*, 1993; Friesen *et al.*, 2001; Fu and Nogaki, 2005). Potential confounding factors such as volume setting, everyday listening mode or gender did not account for this across-listener variation in performance. However, as indicated by the mixed model analysis, across-listener variance was not significant for the

HINT sentences. This might be related to the ceiling effect reported with this measure (Gifford *et al.*, 2008; Faulkner *et al.*, 2001), which might have obscured any significant variation in performance across subjects.

The primary variable manipulated across the two experimental MAPs was site selection based on masked modulation sensitivity; i.e., different electrodes were used in the 10-site MAPs. Certainly, one might argue that differences in performance could also be related to differences in spectral information; the frequency allocation table was the same across the two MAPs, but the individual channel outputs were reallocated to different electrodes in each MAP. However, the fact that better performance on the CUNY test was consistently found for all listeners with the best-MDT MAP weakens this argument and rather suggests that the advantages found for the best-MDT MAP were primarily due to differences in the overall acuity across the two MAPs. Additionally, in support of this argument, some differences were also noted between these two MAPs as described in the following text.

In comparison to the worst-MDT MAP, the best-MDT MAP exhibited better mean MDTs in quiet and in the presence of a masker as well as smaller amounts of masking as expected (see Fig. 6). This suggests that mean temporal acuity was better for the best-MDT MAP than that for the worst-MDT MAP. Alternatively, the reduced modulation sensitivity along with the large amount of masking found for the worst-MDT MAP suggests that the overall acuity of these sites were poorer in comparison to sites in the best-MDT MAP. It is likely that in addition to the reduced temporal processing, these sites were also susceptible to greater overlap in the neural excitation patterns as suggested by the larger amount of masking found for the worst-MDT MAP. Increased overlap in the neural activation has been shown to considerably interfere with the extent to which CI recipients were able to resolve temporal information (Chatterjee, 2003; Richardson *et al.*, 1998) and to also understand speech (Chatterjee and Shannon, 1998; Throckmorton and Collins, 1999; Boëx *et al.*, 2003). Taken together, this perhaps could explain the poor speech performance obtained with the worst-MDT MAP in comparison to the best-MDT MAP.

Based on these findings, one might want to consider using amount of masking instead of masked MDTs to guide the site-selection strategy proposed in this study. However, there are some important considerations that we should bear in mind. For example, given that amount of masking in this study is defined as the absolute difference between masked MDTs and MDTs in quiet, removal of sites with the highest amount of masking does not necessarily lead to removal of those with the poorest MDTs. The results reported earlier in Fig. 3 showed that patterns of MDTs in quiet and in the presence of a masker were strongly correlated, suggesting that temporal acuity was site dependent. Therefore, a site with poor MDTs that were similar with and without a masker would exhibit a smaller amount of masking compared to another site that had good modulation sensitivity in quiet but noticeably elevated MDTs when a masker was present. Figure 2 shows several examples of this scenario. Therefore, the amount of masking derived for each site does not necessarily predict the site's temporal acuity; rather it probably reflects

the spatial properties of that site. The extent to which measures of temporal acuity, spatial acuity or a combination of both are more reflective of subjects' speech recognition needs to be determined to identify the most appropriate tool to capture the functional characteristics of sites of stimulation in multichannel implants. Overall, the findings of this study are consistent with our hypothesis that by eliminating electrodes with worst masked modulation sensitivity, we expect to see better speech recognition performance due to maximizing temporal acuity and minimizing effects of channel interaction.

C. Implications

Results from the current study expand our understanding of across-site variation in the perception of a modulated signal and further shed light on the extent to which temporal acuity is important for speech recognition in electric hearing. Current CI prostheses offer multiple channels that allow different information to be presented to different tonotopic regions. However, the individual contribution of each stimulation site to overall performance does not seem to be equal. It is likely that some of the poorly performing sites are rather detrimental to overall performance. Hence, it is important to identify those poorly performing sites in an attempt to optimize processor fitting strategies.

Site identification and removal based on temporal modulation and masking measures tested in this study holds a promising lead in the clinical domain given that temporal processing in CI users has been shown to account for across-listener variances (Chatterjee and Shannon, 1998; Fu, 2002). In general, the experiments reported here served two purposes (1) to identify a psychophysical measure that is sensitive to across-site variation and (2) to determine the extent to which this measure affects speech performance. The feasibility of using the site-removal approach based on masked modulation sensitivity to guide clinical implant fitting defines the next important phase of this work that is currently being examined in our laboratory.

If the temporal acuity measure used in this work proves to be the most successful tool to identify poorly performing sites, one should also consider other strategies alternative to site-removal to optimize CI users' performance. Given the demonstrated importance of temporal acuity to speech recognition, optimizing temporal processing at those poorly performing sites should, in theory, lead to better speech recognition. The advantages of each approach in relation to speech recognition should certainly be examined in future studies. In addition, the search to find the most appropriate and efficient tool to identify sites with poor electrode-neuron interface should certainly continue.

ACKNOWLEDGMENTS

We are grateful to all of our dedicated implanted subjects for their cheerful participation in this research. We are also grateful to Thyag Sadasivan for programming the psychophysics platforms, to Li Xu and Ning Zhou for providing programs for the speech tests, and Brady West at the University of Michigan Center for Statistical Consultation and Research for consultation on the data analysis. We would

also like to thank Deborah Colesa for helping with subjects' logistics and Chris Ellinger for technical support. This work was supported by NIH/NIDCD Grant Nos. R01 DC010786, P30 DC 05188, and F32 DC010318.

- Bierer, J. A., and Faulkner, K. F. (2010). "Identifying cochlear implant channels with poor electrode-neuron interface: Partial tripolar, single-channel thresholds and psychophysical tuning curves," *Ear Hear.* **31**, 247–258.
- Boëx, C., Kos, M. I., and Pelizzone, M. (2003). "Forward masking in different cochlear implant systems," *J. Acoust. Soc. Am.* **114**, 2058–2065.
- Cazals, Y., Pelizzone, M., Saudan, O., and Boëx, C. (1994). "Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants," *J. Acoust. Soc. Am.* **96**, 2048–2054.
- Chatterjee, M. (2003). "Modulation masking in cochlear implant listeners: Envelope versus tonotopic components," *J. Acoust. Soc. Am.* **113**, 2042–2053.
- Chatterjee, M., and Shannon, R. V. (1998). "Forward masked excitation patterns in multielectrode electrical stimulation," *J. Acoust. Soc. Am.* **103**, 2565–2572.
- Colletti, V., and Shannon, R. V. (2005). "Open set speech perception with auditory brainstem implant," *Laryngoscope* **115**, 1974–1978.
- Eddington, D. K., Dobelle, W. H., Brackman, D. E., Mladejovsky, M. G., and Parkin, J. L. (1978). "Auditory prosthesis research with multiple channel intracochlear stimulation in man," *Ann. Otolaryngol. Suppl.* **53**, 5–39.
- Faulkner, A., Rosen, S., and Wilkinson, L. (2001). "Effects of the number of channels and speech-to-noise ratio on rate of connected discourse tracking through a simulated cochlear implant speech processor," *Ear Hear.* **22**, 431–438.
- Fayad, J. N., Makarem, A. O., and Linthicum, F. H., Jr., (2009). "Histopathological assessment of fibrosis and new bone formation in implanted human temporal bones using 3D reconstruction," *Otolaryngol. Head Neck Surg.* **141**, 247–252.
- Friesen, L. M., Shannon, R. V., Başkent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implant," *J. Acoust. Soc. Am.* **110**, 1150–1163.
- Fu, Q.-J. (2002). "Temporal processing and speech recognition in cochlear implant users," *Neuroreport* **13**, 1635–1639.
- Fu, Q.-J., and Nogaki, G. (2005). "Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing," *J. Assoc. Res. Otolaryngol.* **6**, 19–27.
- Fu, Q.-J., Shannon, R. V., and Wang, X. (1998). "Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing," *J. Acoust. Soc. Am.* **104**, 3586–3596.
- Gantz, B. D., Woodworth, G. G., Abbas, P. J., Knutson, J. F., and Tyler, R. S. (1993). "Multivariate predictors of audiological success with multichannel cochlear implants," *Ann. Otol. Rhinol. Laryngol.* **102**, 909–916.
- Gifford, R. H., Shallop, J. K., and Peterson, A. M. (2008). "Speech recognition materials and ceiling effects: Considerations for cochlear implant programs," *Audiol. Neurotol.* **13**, 193–205.
- 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.
- Hinojosa, R., and Marion, M. (1983). "Histopathology of profound sensorineural deafness," *Ann. N.Y. Acad. Sci.* **405**, 459–484.
- Javel, E., and Shepherd, R. K. (2000). "Electrical stimulation of the auditory nerve. III. Response initiation sites and temporal fine structure," *Hear. Res.* **140**, 45–76.
- Johnsson, L. G., Hawkins, J. E., Jr., Kingsley, T. C., Black, F. O., and Matz, G. J. (1981). "Aminoglycoside-induced cochlear pathology in man," *Acta Otolaryngol. Suppl.* **383**, 1–19.
- Khan, A. M., Whiten, D. M., Nadol, J. B., Jr., and Eddington, D. K. (2005). "Histopathology of human cochlear implants: correlation of psychophysical and anatomical measures," *Hear. Res.* **205**, 83–93.
- Kwon, B., and Turner, C. (2001). "Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference," *J. Acoust. Soc. Am.* **110**, 1130–1140.
- Levitt, H. (1971). "Transformed up-down methods in psychophysics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Luo, X., Fu, Q.-J., Wei, C. G., and Cao, K. L. (2008). "Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users," *Ear Hear.* **29**, 957–970.
- Nadol, J. B., Jr. (1997). "Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation," *Otolaryngol. Head Neck Surg.* **117**, 220–228.
- Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.* **95**, 1085–1099.
- Pfingst, B. E., Burkholder-Juhász, R. A., Xu, L., and Thompson, C. S. (2008). "Across-site pattern of modulation detection in listeners with cochlear implants," *J. Acoust. Soc. Am.* **123**, 1054–1062.
- Qin, M., and Oxenham, A. (2003). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," *Acoust. Soc. Am.* **114**, 446–454.
- Richardson, L., Busby, P., and Clark, G. (1998). "Modulation detection interference in cochlear implant subjects," *J. Acoust. Soc. Am.* **104**, 442–452.
- Shannon, R. V. (1992). "Temporal modulation transfer functions in patients with cochlear implants," *J. Acoust. Soc. Am.* **91**, 2156–2164.
- Shannon, R. V., Jansvold, A., Padilla, M., Robert, M., and Wang, X. (1999). "Consonant recordings for speech testing," *J. Acoust. Soc. Am.* **106**, L71–L74.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Shibata, S. B., Cortez, S. R., Beyer, L. A., Wiler, J. A., Di Polo, A., Pfingst, B. E., and Raphael, Y. (2010). "Transgenic BDNF induces nerve fiber regrowth into the auditory epithelium in deaf cochleae," *Exp. Neurol.* **223**, 464–472.
- Swanson, B. (2004). *Nucleus MATLAB Toolbox 3.02 Software User Manual* (Cochlear Ltd., Lane Cove NSW, Australia).
- Tasaki, I. (1955). "New measurements of the capacity and the resistance of the myelin sheath and the nodal membrane of the isolated frog nerve fiber," *Am. J. Physiol.* **181**, 639–650.
- 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.
- Van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (1987). "Speech waveform envelope cues for consonant recognition," *J. Acoust. Soc. Am.* **82**, 1152–1161.
- Vollmer, M., Snyder, R., Beitel, R., Moore, C., and Rebscher, S. (2000). "Effects of congenital deafness on central auditory processing," in *Proceedings of the 4th European Congress of Oto-Rhino-Laryngology, Berlin*, edited by K. Jahnke and M. Fischer (Monuzzi, Bologna), pp. 181–186.
- Vollmer, M., Snyder, R., Leake, P., Beitel, R., and Rebscher, S. (1999). "Temporal properties of chronic cochlear electrical stimulation determine temporal resolution in neurons in cat inferior colliculus," *J. Neurophysiol.* **82**, 2883–2902.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (1991). "Better speech recognition with cochlear implants," *Nature* **352**, 236–238.
- Xu, L., Thompson, C. S., and Pfingst, B. E. (2005). "Relative contributions of spectral and temporal cues for phoneme recognition," *J. Acoust. Soc. Am.* **117**, 3255–3267.
- Xu, L., Tsai, Y., and Pfingst, B. E. (2002). "Features of stimulation affecting tonal-speech perception: Implications for cochlear prostheses," *J. Acoust. Soc. Am.* **112**, 247–258.
- 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.