Acoustic temporal modulation detection and speech perception in cochlear implant listeners[a]

Jong Ho Won,[b] Ward R. Drennan, Kaibao Nie, Elyse M. Jameyson, and Jay T. Rubinstein
Virginia Merrill Bloedel Hearing Research Center, Department of Otolaryngology-Head and Neck Surgery, University of Washington, Seattle, Washington 98195

(Received 4 June 2010; revised 27 April 2011; accepted 29 April 2011)

The goals of the present study were to measure acoustic temporal modulation transfer functions (TMTFs) in cochlear implant listeners and examine the relationship between modulation detection and speech recognition abilities. The effects of automatic gain control, presentation level and number of channels on modulation detection thresholds (MDTs) were examined using the listeners' clinical sound processor. The general form of the TMTF was low-pass, consistent with previous studies. The operation of automatic gain control had no effect on MDTs when the stimuli were presented at 65 dBA. MDTs were not dependent on the presentation levels (ranging from 50 to 75 dBA) nor on the number of channels. Significant correlations were found between MDTs and speech recognition scores. The rates of decay of the TMTFs were predictive of speech recognition abilities. Spectral-ripple discrimination was evaluated to examine the relationship between temporal and spectral envelope sensitivities. No correlations were found between the two measures, and 56% of the variance in speech recognition was predicted jointly by the two tasks. The present study suggests that temporal modulation detection measured with the sound processor can serve as a useful measure of the ability of clinical sound processing strategies to deliver clinically pertinent temporal information. © *2011 Acoustical Society of America*. [DOI: 10.1121/1.3592521]

PACS number(s): 43.66.Ts, 43.66.Fe, 43.71.Ky [RYL] Pages: 376–388

I. INTRODUCTION

Numerous studies have demonstrated the importance of temporal modulation detection for speech perception outcomes in cochlear implant (CI) users. Amplitude modulation detection and amplitude modulation frequency discrimination are critical for vowel and consonant recognition (Fu, 2002; Cazals *et al.*, 1994), prosodic and segmental speech recognition (Van Tasell *et al.*, 1987), phoneme recognition (Fu and Shannon, 2000; Xu *et al.*, 2005), voice gender recognition (Fu *et al.*, 2004), and lexical-tone recognition (Fu *et al.*, 1998; Fu and Zeng, 2000; Xu *et al.*, 2002; Wei *et al.*, 2007; Luo *et al.*, 2008). Significant correlations were also found between consonant and vowel recognition and temporal modulation detection on single electrodes by Cazals *et al.* (1994), with r^2-values ranged from 0.26 to 0.72, and by Fu (2002), with r^2-values ranged from 0.72 to 0.97.

With single electrodes and direct electric stimulation, a wide range of temporal modulation detection studies have been conducted including the evaluation of temporal modulation transfer functions (TMTFs) (e.g., Shannon, 1992; Busby *et al.*, 1993), the determination of effects of stimulation rate, mode, level (Galvin and Fu, 2005), site (Pfingst *et al.*, 2007, 2008), loudness growth (Galvin and Fu, 2009), envelope masking (Chatterjee and Oba, 2004), and stimulus duration (Luo *et al.*, 2010) on modulation detection thresholds (MDTs), the analysis of the relationship between amplitude modulation detection and intensity discrimination (Donaldson and Viemeister, 2000), and the effect of amplitude modulation on loudness perception (McKay and Henshall, 2010). This approach provides useful information about what electrical amplitude modulation CI users are capable of hearing, and suggest ways future technologies might make best use of these abilities. However, the approach does not assess the subjects' temporal sensitivities using a sound processor with a clinical encoding strategy. If sound processors are used in a modulation detection test, it can reveal (1) what CIs actually provide in practice; and (2) how well different processing approaches deliver critical temporal information, creating a potential tool for clinical and engineering research. CI listeners receive multi-channel stimulation, thus temporal modulation detection ability with multi-channel stimulation might be more pertinent to speech perception abilities.

The primary goal of the present study was to measure MDTs for sinusoidally amplitude-modulated acoustic wideband noise as a function of modulation frequency in CI users with their sound processors. It is hypothesized that temporal modulation transfer functions (TMTFs) will show a lowpass characteristic consistent with previous reports (e.g., Shannon, 1992). Several factors could potentially contribute to the acoustic MDTs such as a front-end automatic gain control (AGC), the presentation level of the stimuli, and use of multiple electrodes. The effect of AGC on MDTs was evaluated acutely by turning off AGC. To examine the effects of acoustic stimulation level on MDTs, three different presentation levels (50, 65, and 75 dBA) were tested. To evaluate the effect of number of channels, MDTs were measured in six

---

[a] This paper was a part of the first author's Ph.D. dissertation. Portions of this work were presented at American Auditory Society meeting, March, 2009, and Conference on Implantable Auditory Prostheses, July, 2009.

[b] Author to whom correspondence should be addressed. Electronic mail: jhwon@uw.edu

0001-4966/2011/130(1)/376/13/$30.00

different channel number conditions: 1, 2, 4, 8, 12, and 16 channels.

Another goal of the present study was to determine the relationship of temporal modulation detection sensitivity with spectral envelope sensitivity and speech perception in CI users. Chatterjee and Yu (2010) showed that modulation detection sensitivity at 100 Hz modulation rate was significantly correlated with electrode discrimination at low levels (20% and 30% of the dynamic range) when subjects were tested in a bipolar mode. When subjects were tested either at higher stimulation levels, at 10 Hz modulation rate, or with a monopolar configuration, modulation detection sensitivity was not correlated with electrode discrimination. They speculated that some common underlying factors, such as the functional health of local populations of neurons, might partially determine subjects' sensitivity to both temporal and spectral measures at specific locations, hence, CI listeners who have good modulation sensitivity would also have good spectral sensitivity, resulting in correlations between the two measures. In the present study, to determine if this relationship between temporal and spectral sensitivity existed when subjects were tested with the sound processors, spectral-ripple discrimination ability was evaluated in the same subjects (Henry *et al.*, 2005; Won *et al.*, 2007). Spectral-ripple discrimination has been shown to be predictive of multiple domains of clinical outcomes in CIs including speech recognition (Henry and Turner, 2003; Henry *et al.*, 2005; Litvak *et al.*, 2007; Saoji *et al.*, 2009) and music perception (Won *et al.*, 2010). Temporal modulation detection evaluates the listener's sensitivity to temporal envelope, while spectral-ripple discrimination evaluates the sensitivity to the spectral envelope of the sound. If the same hearing mechanism is used for the two tests, a significant correlation between the two tests would be observed. Alternatively, if the two tests measure separate hearing abilities, it is expected that the temporal modulation detection and spectral-ripple discrimination will not be correlated with each other, and that the two measures together will correlate with speech perception abilities better than either on its own.

II. EXPERIMENT 1: ACOUSTIC TEMPORAL MODULATION TRANSFER FUNCTIONS IN CI USERS

A. Methods

1. Subjects

Twenty-four post-lingually deafened adult CI users participated in this experiment. Eighteen subjects used Advanced Bionics devices, five subjects used Cochlear devices, and one subject used a MED-EL device. All subjects were native speakers of American English. Table I shows individual subject information. The use of human subjects in the present study was reviewed and approved by the University of Washington Institutional Review Board.

2. Procedure

All subjects listened to the stimuli using their own sound processor set to a comfortable listening level. Five of the 24 subjects were bilateral users (S55, S62, S65, S66, and S71). Four of them were tested with one implant and one subject (S55) was tested with both implants functioning. CI sound processor settings were not changed from the clinical settings in this experiment. All tests were conducted in a double-walled, sound-treated booth (IAC). Custom MATLAB (The Mathworks, Inc.) programs were used to present stimuli on a Macintosh G5 computer with a Crown D45 amplifier. A single loudspeaker (B&W studio monitor DM303 with a frequency and phase response that exceed ANSI standards for speech audiometry), positioned 1-m in front of the subjects, presented stimuli.

The modulation detection test was adapted from Bacon and Viemeister (1985). Acoustic stimuli were 2-s in duration. One of the two 1-s observation intervals consisted of sinusoidally amplitude modulated wide band noise, and the other 1-s observation interval consisted of continuous wide band noise. For the modulated stimuli, sinusoidal amplitude modulation was applied to the wideband noise carrier using the following equation: $[f(t)][1+ m_i \sin(2\pi f_m t)]$, where $f(t)$ is the wideband noise carrier, m_i is the modulation index (i.e., modulation depth), and f_m is the modulation frequency. To compensate for the acoustic intensity increment in the modulated stimuli, the modulated waveform was divided by a factor of $1+(m_i^2/2)$ (Viemeister, 1979; Bacon and Viemeister, 1985). Both the modulated and unmodulated signals were gated on and off with 10 ms linear ramps, then concatenated with no gap between the two signals.

A 2-interval, 2-adaptive forced-choice (AFC) procedure was used to measure the modulation detection thresholds (MDTs). Seven different modulation frequencies were tested including 10, 50, 75, 100, 150, 200, and 300 Hz. Stimuli were presented at 65 dBA. The presentation of stimuli at a conversational level was desirable because the MDTs would be compared to speech recognition. During one of the two 1-s observation intervals, the carrier was sinusoidally amplitude modulated. The subjects were instructed to choose the interval which contains the modulated noise. Visual feedback of the correct answer was given after each presentation. A 2-down, 1-up adaptive procedure was used to measure the modulation depth (m_i) threshold, converging on 70.7% (Levitt, 1971), starting with a modulation depth of 100% and decreasing in steps of 4 dB from the first to the fourth reversal, and 2 dB for the next 10 reversals. For each tracking history, the final 10 reversals were averaged to obtain the MDT for that tracking history. MDTs in dB relative to 100% modulation $[20\log_{10}(m_i)]$ were obtained. Subjects completed all 7 modulation frequencies in random order, and then the subjects repeated a new set of 7 modulation frequencies with newly created random order. Generally, six tracking histories, which generated 60 reversals total, were conducted to determine the average thresholds for each modulation frequency. If subjects could not complete the six tracking histories for some modulation frequencies due to time and scheduling constraints, the mean across the multiple tracking histories (generally 3, 4, or 5) was obtained for the average thresholds. When six tracking histories were done, the modulation detection test required 2–2.5 hours for each subject.

TABLE I. Subject characteristics.

Subject	Age (yr)	Duration of hearing loss (yr)[a]	Duration of implant use (yr)	Etiology	Implant type	Strategy	Experiment participated
S01	61	0.3	2	Unknown	Nucleus 24	ACE	1,4
S03	61	5	11	Genetic	Nucleus 22	SPEAK	1,4
S04	62	1	3	Unknown	Nucleus 24	ACE	1,2,4
S12	49	0	2	Connexin 26	MED-EL Combi40+	CIS	1,2,4
S34	55		1.5	Noise exposure	HiRes90K	HiResolution	1,2,3,4
S38	51	9	4	Noise exposure	Nucleus 24	ACE	1,4
S40	72	5	6	Genetic	HiRes90K	HiResolution	1,4
S41	52	7	5	Hereditary	HiRes90K	HiResolution	1,4
S48	67	10	0.5	Unknown	HiRes90K	HiResolution	1,2,3,4
S49	64	4	0.75	Hereditary	HiRes90K	Fidelity120	1,4
S51	56	7	6	Hereditary	Clarion CII	HiResolution	1,4
S52	77	0	0.5	Noise exposure	HiRes90K	Fidelity120	1,2,3
S53	63	3	7	Unknown	Clarion CII	Fidelity120	1,4
S54	25	0.5	2.5	Unknown	HiRes90K	HiResolution	1,4
S55	65	40	1	Genetic	HiRes90K	Fidelity120	1,4
S58	64	57	7	Noise exposure	Clarion CII	Fidelity120	1,4
S59	47	12	2.5	Noise exposure	HiRes90K	Fidelity120	1,4
S61	78	10	1	Genetic	HiRes90K	Fidelity120	1,4
S62	32	3	1	Unknown	HiRes90K	Fidelity120	1,4
S65	56	2	7	Unknown	Clarion CII	HiResolution	1,4
S66	66	3	2	Unknown	HiRes90K	Fidelity120	1,4
S69	60	30	2	Unknown	HiRes90K	Fidelity120	1,2,4
S70	59	0	1	Genetic	Freedom	ACE	1,4
S71	70	15	1.5	Genetic	HiRes90K	Fidelity120	1,2,3,4

[a] The duration of their hearing loss before implantation

B. Results

Modulation detection sensitivity decreased as the modulation frequency was increased. The left panel of Fig. 1 shows the acoustic TMTFs for the 24 CI subjects, plotting the MDTs as a function of modulation frequency. The right panel of Fig. 1 shows the average TMTFs for the 24 CI subjects. The overall pattern of the acoustic TMTFs for the CI subjects shows low-pass filter characteristics, which is consistent with the previously reported TMTFs in CI listeners measured with single-electrode, direct stimulation (e.g., Shannon, 1992).

For comparison, TMTFs for normal-hearing (NH) and for sensorineural hearing-impaired (HI) listeners are also shown (Bacon and Viemeister, 1985) in Fig. 2. Those NH and HI listeners (Figs. 1 and 2 in Bacon and Viemeister, 1985) were tested with the same method in this study except that the spectrum level of the modulated noise measured at 1 kHz was 30 dB, which corresponds to an overall level of 77 dB sound pressure level measured in a 6-cc coupler (Viemeister, 1979; Bacon and Viemeister, 1985). For all modulation frequencies, the CI subjects showed considerably worse MDTs than NH or HI listeners. Though similar low-pass patterns of the functions are observed in the three different subject groups, the slope of the functions was steepest in the CI subjects. It is also apparent that the difference between the MDTs in the CI subjects and the NH/HI listeners becomes greater as the modulation frequency increases.

A 7 × 6 repeated measures analysis of variance (7 modulation frequencies and 6 repetitions) was performed on the

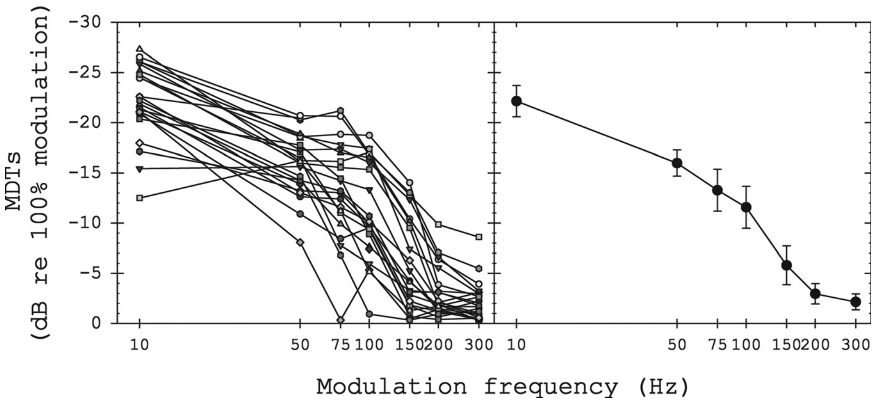

FIG. 1. The left panel shows the individual TMTFs for 24 CI subjects. The right panel shows the average TMTFs for the 24 CI subjects. Error bars indicate 95% confidence intervals across the subjects for modulation detection thresholds at each modulation frequency.

MDTs measured in 12 subjects who completed 6 repetitions for all 7 modulation frequencies. It showed a significant effect of modulation frequency ($F_{6,66} = 168.2$, $p < 0.0001$) and repetition ($F_{5,55} = 6.4$, $p < 0.0001$). There was no interaction between the two ($p = 0.2$). A separate one-way analysis of variance on the MDTs at 7 modulation frequencies in all 24 subjects showed a significant effect of modulation frequency ($F_{6,154} = 86.3$, $p < 0.0001$). A *post hoc* analysis using the Tukey test was performed to determine if any of MDTs are different from each other for each modulation frequency. MDTs were generally significantly different ($p < 0.01$) from each other except for MDTs between 1 or 2 neighboring modulation frequencies. MDTs at 10 Hz were significantly different from all of other frequencies ($p < 0.001$).

Considerable variability was observed in the patterns of TMTFs across the 24 subjects. To better characterize the pattern and shape of each modulation transfer function, MDTs were fitted with an exponential function with two parameters using the following equation:

$$|f(x)| = Ae^{bx}, \tag{1}$$

where $|f(x)|$ is the absolute value of the MDT in dB, x is the modulation frequency. The rate of the exponent (b) and the y intercept (A) were assessed. The exponential fit accounted for an average of 86% of the variance in the TMTFs for each of the 24 subjects. To quantify the difference among NH, HI, and CI subjects shown in Fig. 2, the NH and HI data were also fitted with the exponential function. The average exponent (b) for NH, HI, and CI subjects were −0.0015, −0.003, and −0.01, demonstrating that CI subjects showed the steepest slope, and HI subjects showed a steeper slope than NH subjects.

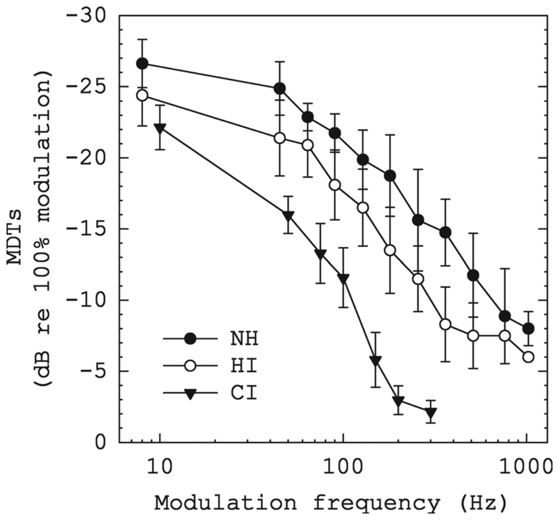

FIG. 2. Comparison of the average TMTFs measured in 24 CI listeners in this study (filled circles), 4 normal-hearing (NH, open circles) listeners, and in 5 hearing-impaired (HI, filled triangles) listeners. The NH and HI data are taken from Bacon and Viemeister (1985), which measured the TMTFs with a continuous wide-band noise carrier. The spectrum level of the noise was 30 dB (measured at 1 kHz), which corresponds to an overall level of 77 dB sound pressure level. Error bars indicate 95% confidence intervals across the subjects for modulation detection thresholds at each modulation frequency.

III. EXPERIMENT 2: EFFECT OF PRESENTATION LEVELS AND AUTOMATIC GAIN CONTROL ON MODULATION DETECTION THRESHOLDS

A. Rationale

Previous studies (e.g., Fu, 2002; Galvin and Fu, 2005, 2009) demonstrated that MDTs determined with single-electrode, direct stimulation are dependent on stimulation levels, showing better detection ability with high stimulation levels. To determine if similar level effects occur in the MDTs measured acoustically using the sound processor, three different presentation levels (50, 65, and 75 dBA) were tested.

The effect of AGC on MDTs was also evaluated in Experiment 2. AGC is operated in the clinical processors, so it is possible that the nonlinear characteristic of AGC could affect MDTs. For example, if instantaneous levels are too high, AGC could reduce the depth of modulation and thus increase MDTs (i.e., worse performance) measured through the processor. The AGC effect was tested by comparing the MDTs with and without the operation of AGC. If AGC has no effect on the results, it will suggest it is inconsequential to the modulation detection test.

B. Methods

1. Subjects

Seven subjects (5 Advanced Bionics device users, 1 MED-EL device user, and 1 Nucleus device user) were tested with their sound processors in an experiment to determine the effect of presentation levels on MDTs. The five Advanced Bionics users were also tested in an experiment to examine the effect of AGC on MDTs with a laboratory-owned Platinum Series Sound Processor.

2. Procedure

In order to examine the effect of presentation levels, three different levels were tested including 50, 65, and 75 dBA at three different modulation frequencies (10, 100, and 300 Hz). The procedure for determining MDTs was the same as with Experiment 1 except that three tracking histories for each testing condition were used in this experiment. The subjects were instructed not to change the volume control of the sound processor during the testing. The order of presentation was randomized across subjects for the level and modulation frequency.

The effect of AGC on MDTs was also evaluated acutely by turning off AGC in the Advanced Bionics devices using the SoundWave clinical fitting software. Modulation frequencies of 10, 100, and 300 Hz were tested. The presentation level was 65 dBA to determine the extent to which the AGC contributed to the MDTs measured in Experiment 1. The same modulation detection test was used. Other map parameters such as T (threshold) and M (maximum comfortable loudness) levels, pulse rate, pulse width, and input dynamic range were left unchanged. The Advanced Bionics device uses a dual-level AGC system (Moore *et al.*, 1991; Firszt, 2003). The slow-acting control has a compression threshold of 57 dB sound pressure level (325 ms attack and

1000 ms release times) and the fast-acting control has a higher compression threshold of 65 dB sound pressure level (attack time is less than 0.6 ms and 8 ms release times). Thus, the AGC was operated during the testing in Experiment 1 because the stimuli were presented at 65 dBA. The order of presentation was randomized across subjects for the AGC condition (i.e., on and off) and modulation frequency.

In addition to the behavioral testing, quantitative analyses were done with the electrode outputs to further analyze the effect of AGC and presentation level. The Platinum Series body-worn processor of Advanced Bionics was used for measurement using a single subject's HiResolution program. The processor was positioned 1-m in front of the loudspeaker. Stimuli were sinusoidally amplitude modulated (SAM) wide band noise of 1-s duration. A modulation frequency of 10 Hz was used. To evaluate the effect of AGC on the modulation pattern in the electrode outputs, SAM noise with three different modulation depths (0, −10, and −20 dB relative to 100% modulation) was presented. To evaluate the effect of presentation level and its interaction with the operation of AGC, four different levels were presented including 50, 65, 75, and 85 dBA. For this measurement, a modulation frequency of 10 Hz and 0 dB modulation depth (i.e., 100% modulation) was used. For all the measurements, current values (mA) at modulation peaks and valleys, peak-to-valley ratios, and root-mean-square values of resulting current pulses were analyzed.

All of the measurements were made in a double-walled, sound-treated booth (IAC). The current stimulation produced on an electrode was measured using an "implant-in-a-box," which is composed of a CI receiver/stimulator with electrode outputs linked to 5-kΩ load-resistors. Three different electrodes were used for the measurements including electrode 1, 8, and 16. Responses were measured and sampled at a rate of 25 kHz using a Tektronix DPO 4034 Digital Phosphor Oscilloscope (Beaverton, OR).

C. Results

MDTs at 10, 100, and 300 Hz showed no presentation level effect. The left panel of Fig. 3 shows the mean MDTs as a function of modulation frequency. A 3 × 3 × 3 analysis of variance (3 levels, 3 modulation frequencies, 3 repetitions) showed that there was no effect of level ($F_{2,12} = 0.34$, $p = 0.72$) and no effect of repetition ($F_{2,12} = 1.37$, $p = 0.29$). As expected, modulation frequency had a significant effect ($F_{2,12} = 100.37$, $p < 0.0001$).

AGC also did not have any effect on performance. The right panel of Fig. 3 shows the mean MDTs with AGC-on and AGC-off. A 2 × 3 × 3 analysis of variance (AGC on and off, 3 modulation frequencies, 3 repetitions) showed that there was no effect of AGC ($F_{1,3} = 0.77$, $p = 0.44$), no effect of repetition ($F_{2,6} = 0.42$, $p = 0.67$), and a significant effect of modulation frequency ($F_{2,6} = 28.52$, $p < 0.001$).

Figure 4 shows the resulting electrode outputs for electrode 1 (the lowest frequency channel) in response to the input stimuli of SAM noise at 10 Hz with a modulation depth of 0 dB, −10 dB, and −20 dB relative to 100% modulation (i.e., $20\log_{10}(m_i)$). The presentation level was 65 dBA. The left and right panels of Fig. 4 show the electrode outputs with the AGC on and off conditions, respectively, and do not show any distinguishable differences. Figure 5 shows the resulting electrode outputs in response to the input stimuli of SAM noise at 10 Hz with 0 dB modulation depth (i.e., 100% modulation). Four different levels were presented with and without the operation of AGC. At moderate presentation levels (50–65 dBA), the modulation patterns in the electrode outputs were comparable.

Table II shows the mean current values (mA) for 10 peaks and 10 valleys over 1-s and their average peak-to-valley ratios. Table III shows root-mean-square values of current pulses for the electrode outputs. Overall, the current output values were greater with the AGC off condition, but the difference was minimal at 65 dBA, which was used to determine the acoustic TMTF in Experiment 1 (Fig. 1) and to evaluate the effect of AGC on MDTs at three modulation frequencies (the right panel of Fig. 3). When the stimuli were presented at soft to conversational level (50–65 dBA), similar peak-to-valley ratios were observed, meaning that about the same amount of modulation was presented with and without the operation of AGC. But when the level was very high (85 dBA), the operation of the AGC increased the peak-to-valley ratio, thus it provided a greater amount of modulation in the outputs. This is also reflected in the

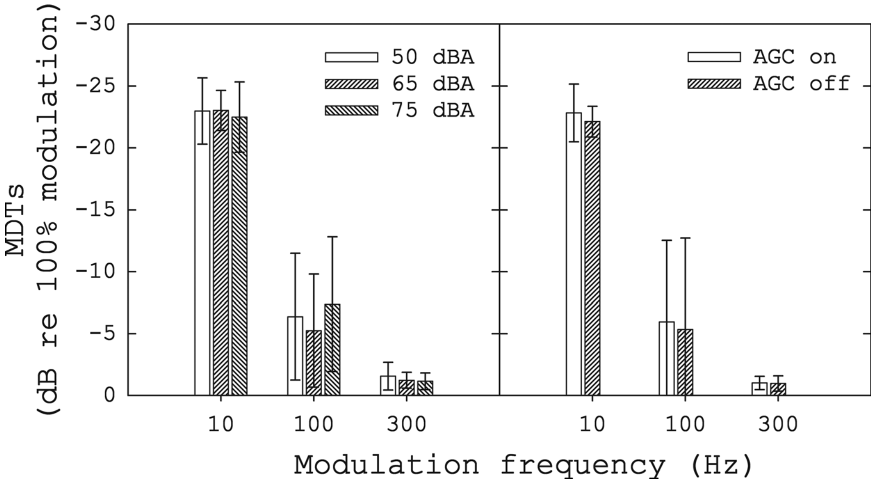

FIG. 3. The left panel shows modulation detection thresholds for three different presentation levels in 7 subjects. The right panel shows modulation detection thresholds measured with and without the operation of automatic gain control in 5 subjects. The presentation level was 65 dBA. Error bars represent ±95% confidence intervals of the mean across subjects.

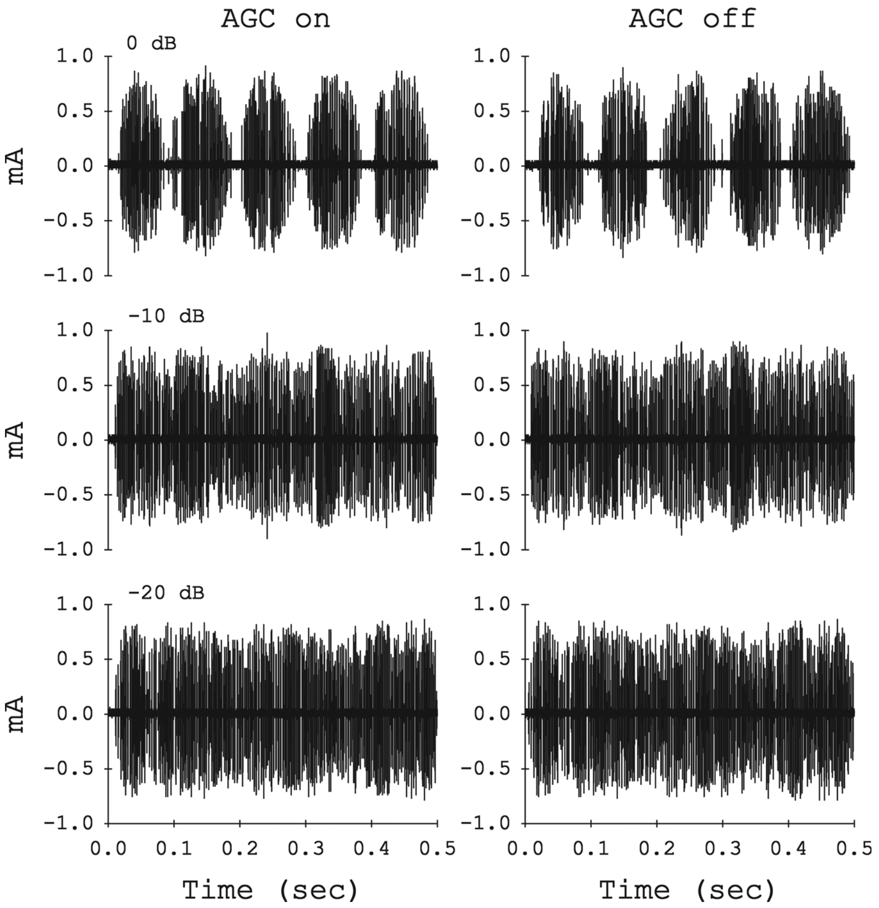

FIG. 4. The electrode outputs in response to different amount of modulation depth of the sinusoidally amplitude modulated (SAM) noise at 10 Hz. Advanced Bionics Platinum sound processor was used with HiResolution processing strategy. Electrode 1 (the lowest frequency channel) was used. The left panel shows the electrode outputs recorded with the operation of AGC. The right panel shows the electrode outputs recorded with the absence of the AGC. Each row shows the electrode outputs in response to the SAM noise with a modulation depth of 0 dB, −10 dB, and −20 dB, respectively. The first 0.5 s are shown.

electrode outputs for electrode 16 (the highest frequency channel) in Fig. 5. It appears that AGC would have a beneficial effect in representing the modulation pattern at higher acoustic levels (75–85 dBA). The electrode output analyses (Table II, Figs. 4 and 5) also show that the effects of AGC vary across channels. The acoustic-input/electric-output relationship is not linear and there are several factors that could affect the relationship between the two such as the time constant and compression function of the AGC, maximum levels for target electrodes, pulse width, and input dynamic range. For electrode 1 with AGC on, the maximum current values for 50, 65, and 75 dBA input stimuli were about 48%, 74%, and 83% of the subject's dynamic range (1.26 mA) of that electrode.

IV. EXPERIMENT 3: THE EFFECT OF NUMBER OF CHANNELS ON MODULATION DETECTION THRESHOLDS

A. Rationale

Another primary difference between the present study and previous single-electrode, direct stimulation studies is the use of multiple electrodes. With more than one electrode and direct stimulation, previous studies investigated modulation masking pattern (Chatterjee, 2003), envelope interaction (Chatterjee and Oba, 2004), modulation detection interference (Richardson *et al.*, 1998), and the effect of multiple electrodes on pulse rate discrimination (McKay *et al.*, 2005; Carlyon *et al.*, 2010), but the effect of number of channels on MDTs in a sound processor context has not, to the authors' knowledge, been previously reported. In this experiment, when multiple electrodes are used, the filter bandwidth for each channel proportionally decreases as the number of channels increases. Eddins (1993) showed that in normal-hearing listeners, modulation detection performance decreased with decreasing stimulus bandwidth, thus modulation detection performance could decrease in CI listeners when more channels are driven by narrower bandwidths. On the other hand, when multiple electrodes are used, modulation detection performance could increase because a wider portion of the auditory-nerve fiber array is excited. In this experiment, the effect of number of channels on MDTs was evaluated by testing 6 different channel conditions.

B. Methods

1. Subjects

Four Advanced Bionics device users participated in this experiment.

2. Procedure

The same procedure for determining MDTs described in Experiment 1 was used except that only one modulation frequency (50 Hz) was tested. The modulation frequency of 50 Hz was chosen because this frequency was frequently used in previous single-electrode, direct stimulation studies (e.g.,

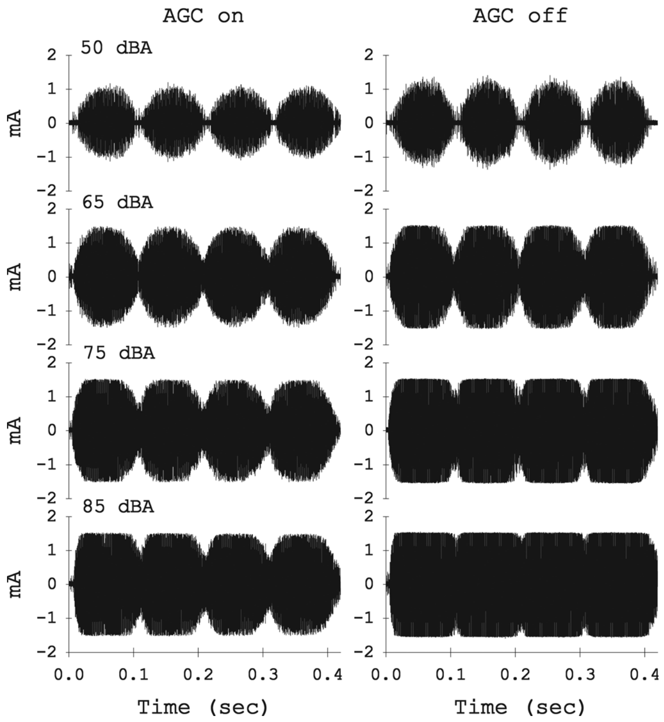

FIG. 5. The electrode outputs in response to different presentation level of the sinusoidally amplitude modulated (SAM) noise at 10 Hz. Advanced Bionics Platinum sound processor was used with the HiResolution processing strategy. Electrode 16 (the highest frequency channel) was used. The left panel shows the electrode outputs recorded with the operation of AGC. The right panel shows the electrode outputs recorded with the absence of the AGC. Each row shows the electrode outputs in response to the SAM noise presented at 50, 65, 75, and 85 dBA, respectively. The first 0.5 s are shown.

Cazals *et al.*, 1994; Chatterjee, 2003; Chatterjee and Oba, 2004; Pfingst *et al.*, 2007; Galvin and Fu, 2009). This modulation frequency was also desirable because 50-Hz is less than any of bandwidths for 16 bands when 16 channels were active. The MDTs at 50 Hz in Experiment 1 did not also show a ceiling or floor effect. Six different maps with the HiResolution strategy were created on a laboratory-owned Platinum sound processor including 1, 2, 4, 8, 12, and 16 channels. The filter cutoff frequencies were determined by the default settings used in the HiResolution strategy for each of these channel conditions, thus the filter bandwidths for individual electrodes increased as the number of channels decreased. The total bandwidth of 250–8700 Hz for the Platinum sound processor remained the same. The electrodes enabled for each of the conditions are listed in Table IV.

C. Results

Figure 6 shows that temporal modulation detection did not change as the number of channels varied. A 6 × 6 analysis of variance (6 different channel conditions, 6 repetitions) demonstrated that there was no effect of the number of channels ($F_{5,10} = 1.24$, $p = 0.36$) and no effect of repetition ($F_{5,10} = 1.55$, $p = 0.26$). A separate one-way analysis of variance also showed no effect of number of channels ($p = 0.91$).

V. EXPERIMENT 4: RELATIONSHIP OF TEMPORAL MODULATION DETECTION WITH SPECTRAL-RIPPLE DISCRIMINATION AND SPEECH PERCEPTION

A. Rationale

Previous single-electrode, direct stimulation studies have shown that MDTs were significantly correlated with vowel and consonant recognition abilities in CI users (e.g., Cazals *et al.*, 1994; Fu, 2002; Luo *et al.*, 2008). To determine if the acoustic MDTs are also predictive of speech perception abilities in CI users, the acoustic MDTs were compared to consonant-nucleus-consonant (CNC) words (Peterson and Lehiste, 1962) recognition scores and speech reception thresholds (SRTs) in steady-state, speech-shaped noise (Turner *et al.*, 2004; Won *et al.*, 2007).

In addition, temporal modulation detection abilities were compared to spectral-ripple discrimination abilities. A recent study by Chatterjee and Yu (2010) showed that the temporal modulation detection at 100 Hz is significantly correlated with electrode discrimination at a particular electrode location. In this experiment, spectral-ripple discrimination was measured to assess CI listener's ability to detect and discriminate spectral envelope changes across channels (Won *et al.*, 2007). It is important to determine if the same relationship exists between temporal modulation detection and spectral-ripple discrimination when sound processors and multiple electrodes are used in the tests, and furthermore, how the two tests together relate to speech perception abilities in CI users.

B. Methods

1. Subjects

The twenty-four subjects in Experiment 1 participated in spectral-ripple discrimination, CNC word recognition, and spondee word recognition in noise tests.

TABLE II. Electrode outputs in response to the sinusoidal amplitude modulated noise at 10 Hz modulation frequency. The mean current values (mA) for 10 peaks (first value) and 10 valleys (second value) over 1-s are shown. The third value (***bold italic***) shows average peak-to-valley ratios. A greater peak-to-valley ratio means a greater amount of modulation is represented in the outputs.

	Electrode 1		Electrode 8		Electrode 16	
Level	AGC on	AGC off	AGC on	AGC off	AGC on	AGC off
50 dBA	0.55, 0.1, ***5.5***	0.66, 0.1, ***6.6***	0.74, 0.1, ***7.4***	0.96,0.1, ***9.6***	1.16, 0.1, ***11.6***	1.34, 0.11, ***12.2***
65 dBA	0.87, 0.1, ***8.7***	1.05, 0.1, ***10.5***	1.05, 0.1, ***10.5***	1.32, 0.15, ***8.8***	1.48, 0.5, ***2.96***	1.52, 0.61, ***2.5***
75 dBA	0.93, 0.1, ***9.3***	1.27, 0.1, ***12.7***	1.1, 0.1, ***11***	1.47, 0.32, ***4.6***	1.49, 0.56, ***2.7***	1.54, 0.94, ***1.6***
85 dBA	0.93, 0.1, ***9.3***	1.33, 0.65, ***2.1***	1.13, 0.24, ***4.7***	1.49, 0.66, ***2.3***	1.5, 0.75, ***2***	1.55, 1.27, ***1.2***

TABLE III. Root-mean-square current values (mA) over 1-s are shown.

	Electrode 1		Electrode 8		Electrode 16	
Level	AGC on	AGC off	AGC on	AGC off	AGC on	AGC off
50 dBA	0.12	0.15	0.32	0.40	0.58	0.73
65 dBA	0.28	0.36	0.55	0.69	0.90	1.10
75 dBA	0.32	0.52	0.58	0.87	0.95	1.21
85 dBA	0.33	0.60	0.62	0.97	1.00	1.27

2. *Procedure*

a. Spectral-ripple stimuli discrimination test. The spectral-ripple discrimination test in this study is the same as that previously described by Won *et al.* (2007). Two-hundred pure-tone frequency components were summed to generate the rippled noise stimuli. The 200 tones were spaced equally on a logarithmic frequency scale. The amplitudes of the components were determined by a full-wave rectified sinusoidal envelope on a logarithmic amplitude scale. The ripple peaks were spaced equally on a logarithmic frequency scale. The stimuli had a bandwidth of 100–5000 Hz and a peak-to-valley ratio of 30 dB. The mean presentation level of the stimuli was 61 dBA and randomly roved in 7 dB range in 1-dB steps. The starting phases of the components were randomized for each presentation. The ripple stimuli were generated with 14 different densities, measured in ripples per octave. The ripple densities differed by ratios of 1.414 (0.125, 0.176, 0.250, 0.354, 0.500, 0.707, 1.000, 1.414, 2.000, 2.828, 4.000, 5.657, 8.000, and 11.314 ripples/octave). Standard (reference stimulus) and inverted (ripple phase reversed test stimulus) ripple stimuli were generated. For standard ripples, the phase of the full-wave rectified sinusoidal spectral envelope was set to zero radians, and for inverted ripples, it was set to $\pi/2$. The stimuli had 500 ms total duration and were ramped with 150 ms rise/fall times. Stimuli were filtered with a long-term, speech-shaped filter. A 3-AFC, two-up and one-down adaptive procedure was used to determine the spectral-ripple resolution threshold converging on 70.7% correct (Levitt, 1971). Each adaptive track started with 0.176 ripples/octave and moved in equal ratio steps of 1.414. Feedback was not provided. The threshold for a single adaptive track was estimated by averaging the ripple spacing (the number of ripples/octave) for the final 8 of 13 reversals. Six adaptive tracks were repeated to determine the average thresholds.

b. Speech reception threshold (SRT) in steady-state noise. In this SRT in noise test, the subjects were asked to identify one randomly chosen spondee word out of a closed-set of 12 equally difficult spondees (Harris, 1991) in the presence of speech-shaped, steady-state noise (Turner *et al.*, 2004; Won *et al.*, 2007). The spondees, two-syllable words with equal emphasis on each syllable (e.g., "birthday," "padlock," "sidewalk"), were recorded by a female talker (F0 range: 212–250 Hz, Turner *et al.*, 2004). Duration of the steady-state noise was 2.0-s and the onset of the spondees was 500 ms after the onset of the noise. A closed-set, 12-AFC task with 1-up, 1-down adaptive tracking procedure was used to determine speech reception thresholds (SRTs), converging on 50% correct (Levitt, 1971). The level of the target speech was fixed at 65 dBA. The noise level was varied with a step size of 2-dB. Feedback was not provided. For

TABLE IV. Lower and upper cutoff frequencies (Hz) for each channel condition are shown. The active electrode used for each channel is shown in parentheses.

	Frequency allocation for each channel condition					
Number of channels	1	2	4	8	12	16
1	250–8700 (2)	250–1387 (2)	250–698 (4)	250–494 (2)	250–440 (1)	250–416 (1)
2		1389–8700 (8)	698–1386 (8)	494–698 (4)	440–554 (2)	416–494 (2)
3			1386–2762 (12)	698–982 (6)	554–698 (3)	494–587 (3)
4			2762–8700 (16)	982–1388 (8)	698–876 (5)	587–697 (4)
5				1388–1956 (10)	876–1104 (6)	697–828 (5)
6				1956–2764 (12)	1104–1386 (7)	828–983 (6)
7				2764–3896 (14)	1386–1746 (9)	983–1168 (7)
8				3896–8700 (16)	1746–2194 (10)	1168–1387 (8)
9					2194–2746 (11)	1387–1648 (9)
10					2746–3472 (13)	1648–1958 (10)
11					3472–4376 (14)	1958–2326 (11)
12					4376–8700 (15)	2326–2762 (12)
13						2762–3281 (13)
14						3281–3898 (14)
15						3898–4630 (15)
16						4630–8700 (16)

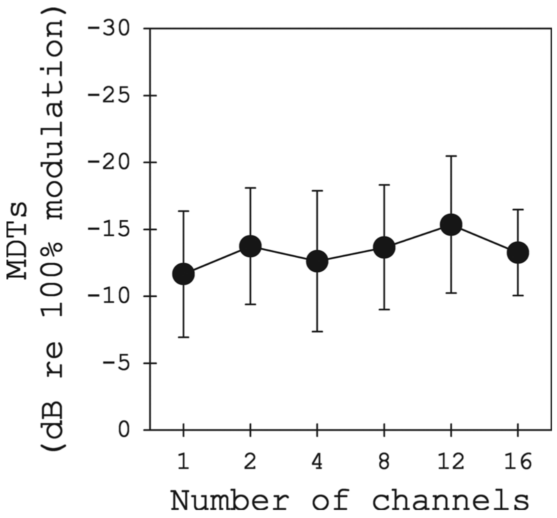

FIG. 6. Modulation detection thresholds as a function of the number of channels. Error bars represent ±1 standard deviation of the mean across subjects.

all subjects, the adaptive track started with +10 dB signal-to-noise ratio (SNR) condition. The threshold for a single adaptive track was estimated by averaging the SNR for the final 10 of 14 reversals. Six adaptive tracks were repeated to determine the average thresholds.

c. CNC word recognition test. Fifty CNC monosyllabic words (Peterson and Lehiste, 1962) were presented in quiet at 62 dBA. A CNC word list was randomly chosen out of ten lists for each subject. The subjects were instructed to repeat the word that they heard. A total percent correct score was calculated after 50 presentations as the percent of words correctly repeated.

C. Results

As shown in Fig. 1, there was considerable variability in the MDTs across the subjects for each modulation frequency. To determine whether there is a relationship between MDTs and speech recognition scores, correlation analyses were performed. Bonferroni corrections were used for all analyses. The left panel of Fig. 7 shows that mean MDTs, averaged across the 7 modulation frequencies, were significantly correlated with CNC word recognition scores ($r = -0.65$, $p = 0.007$). The right panel of Fig. 7 shows the scattergram of mean MDTs and SRTs in noise ($r = 0.57$, $p = 0.03$). In the present study, one of the 24 subjects was tested with bilateral implants. When this subject was excluded from the correlation analyses, roughly the same correlations were observed.

Table V shows the correlations of MDTs at each modulation frequency with CNC word recognition scores and SRTs in steady-state noise. Significant correlations were also found between the exponent b in the exponential fit [Eq. (1)] and CNC word scores ($r = 0.53$, $p = 0.008$) and SRTs in noise ($r = -0.58$, $p = 0.003$). The exponent b controls the rate of decay of the function. In contrast, correlations between the scalar A (i.e., y intercept) and both speech measures were not significant ($p > 0.3$). Thus, temporal modulation detection ability measured with the subjects' sound processors was associated with better speech recognition ability in CI listeners.

To determine if spectral envelope sensitivity was associated with temporal modulation detection, spectral-ripple discrimination thresholds were compared to MDTs. None of the MDTs at the 7 modulation frequencies were correlated with spectral-ripple discrimination thresholds (r-values ranged from -0.07 to -0.42, $p > 0.3$, $N = 24$), suggesting that temporal modulation detection and spectral-ripple discrimination independently assessed different acoustic sensitivities. Significant correlations were found between spectral-ripple discrimination and CNC word recognition scores ($r = 0.61$, $p = 0.002$) and SRTs in steady-steady noise ($r = -0.58$, $p = 0.003$), consistent with a previous report (Won *et al.*, 2007).

Multiple-factor linear regression analyses were performed to determine if the combination of MDTs and spectral-ripple thresholds could predict speech perception performance better than the modulation detection test or spectral-ripple test alone. To be considered valid, the adjusted R^2 for the combination of factors had to be greater than the R^2 of the independent variables, the p-value for each coefficient had to be less than 0.05, the p-value for the F-ratio had to be less than 0.05, the 95% confidence interval of each of the

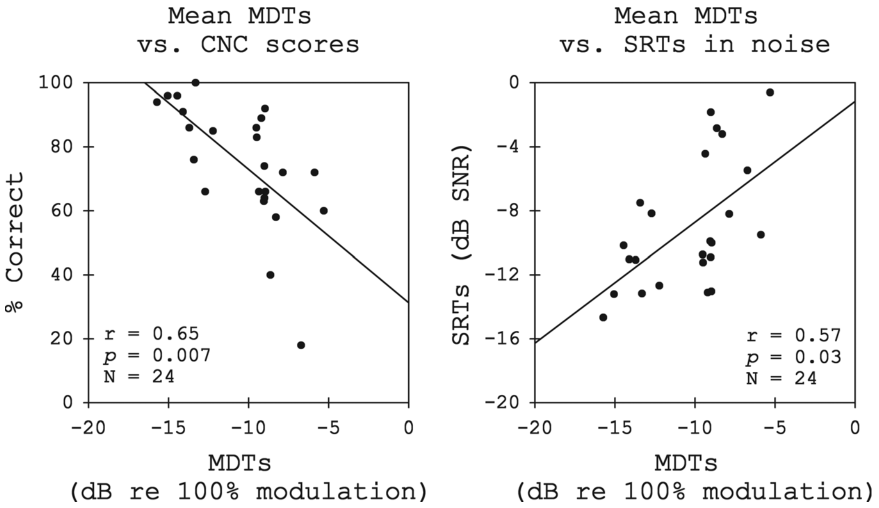

FIG. 7. Relationship between temporal modulation detection and CNC word recognition (left) and speech recognition in steady-state noise (right). Linear regression is represented by the solid line.

TABLE V. Correlations of modulation detection thresholds with CNC word recognition scores and speech reception thresholds (SRTs) in steady-noise.

	10 Hz	50 Hz	75 Hz	100 Hz	150 Hz	200 Hz	300 Hz
CNC scores	−0.25	−0.33	−0.63[a]	−0.63[b]	−0.69[b]	−0.60[a]	−0.58[a]
SRTs in noise	0.36	0.41	0.54	0.41	0.56[a]	0.56[a]	0.59[a]

[a]Significant at $p < 0.05$.

[b]Significant at $p < 0.01$.

regression coefficients could not include 0, and the regression standardized residual had to be normally distributed (Lomax, 2005). Table VI shows the multiple linear regression data which met the criteria described above. These multiple linear regression analyses showed that when coupling temporal modulation detection with spectral ripple discrimination, spectral-ripple discrimination ability accounted for *an additional* 13% to 25% of the variance of the speech performance. Figure 8 shows that 56% of the variance ($r = 0.78$) in CNC word recognition scores and 45% of the variance ($r = 0.71$) in SRTs in steady-noise was accounted for by the combination of mean MDTs and spectral-ripple thresholds.

The 24 CI subjects in this study used different sound processor settings including different encoding strategies. To determine if MDTs obtained with a specific sound processing strategy could predict speech perception performance, correlations were analyzed in subsets of subjects using Fidelity120 and HiResolution strategies. In 11 Fidelity120 users, a correlation of 0.75 ($p = 0.008$) was found between mean MDTs and SRTs in steady-state noise and a correlation of −0.74 ($p = 0.009$) was found between mean MDTs and CNC scores. For 7 HiResolution users, the correlations were 0.63 ($p > 0.1$) and −0.73 ($p = 0.06$) for SRTs and CNC scores, respectively. The correlations for the 7 HiResolution users were not significant at $p = 0.05$, but the range of 95% confidence intervals for correlation in the 24 subjects overlap with them, suggesting that the non-significance may have been due to smaller sample size.

VI. DISCUSSION

A. Summary of results

Experiment 1 demonstrated that the form of the acoustic TMTFs reported in the present study is similar to low-pass shape TMTFs measured by previous investigators (e.g., Shannon, 1992; Busby *et al.*, 1993) using single electrodes. Experiment 2 showed that MDTs were not dependent on the operation of AGC or the presentation level within the range 50–75 dBA. Experiment 3 showed that MDTs were not changed as the number of channels and their filter bandwidths varied together. Experiment 4 showed that MDTs were significantly correlated with speech understanding ability, which is consistent with the previous reports using single electrodes (Cazals *et al.*, 1994; Fu, 2002). These observations suggest that the temporal modulation detection test using the clinical sound processor provides clinically relevant information about temporal modulation sensitivity in CI listeners.

B. Effect of presentation levels and AGC

In the present study, the sinusoidally modulated waveforms were attenuated by a factor of $(1 + m^2/2)$, where m is the modulation index, so that the average acoustic intensities of the modulated and unmodulated waveforms were equal, as done by Viemeister (1979) and Bacon and Viemeister (1985). In order to determine if any acoustic intensity cue affected modulation detection thresholds, we tested two subjects with a level rove (6 dB range with 1 dB steps). An inter-stimulus interval of 300 ms was added between the two observation intervals. Modulation frequencies of 10, 100, and 300 Hz were tested. The two subjects showed no difference in performance between roving and non-roving conditions. This result suggests that subjects performed the modulation detection test by discriminating modulation cues, not an intensity cue.

Unlike previous studies (e.g., Fu, 2002; Galvin and Fu, 2005), the 7 CI subjects in the current study did not show any difference in MDTs across the three presentation levels. It is not clear why the level effect was not seen in the present study. There were some subjects in Fu (2002) and Galvin and Fu (2005) who did not show a level effect when the stimulation level was increased above 50% of dynamic range. In the present study, in the case of electrode 1, the maximum current value for 50 dBA input stimuli was already almost 50% of the subject's dynamic range, and for 65 and 75 dBA input stimuli, they were above 70% of the dynamic range. Therefore, it is possible that the level effect on modulation detection could be immeasurably small at these levels.

TABLE VI. Multiple linear regression analysis results for combination of temporal modulation detection thresholds (MDTs) and spectral-ripple thresholds. Note that no correlations were found between modulation detection thresholds and spectral-ripple discrimination thresholds.

Independent variables		Dependent variable	R^2 (MDTs)	R^2 (Ripple thresholds)	Adjusted R^2 (multiple regression)
MDTs at 75 Hz	Spectral-ripple thresholds	CNC scores	0.40	0.37	0.56
MDTs at 100 Hz		CNC scores	0.40	0.37	0.58
MDTs at 200 Hz		CNC scores	0.36	0.37	0.51
MDTs at 300 Hz		CNC scores	0.34	0.37	0.53
Mean MDTs		CNC scores	0.42	0.37	0.56
MDTs at 50 Hz		SRTs in noise	0.17	0.34	0.42
MDTs at 75 Hz		SRTs in noise	0.29	0.34	0.44
MDTs at 200 Hz		SRTs in noise	0.31	0.34	0.44
MDTs at 300 Hz		SRTs in noise	0.35	0.34	0.51
Mean MDTs		SRTs in noise	0.32	0.34	0.45

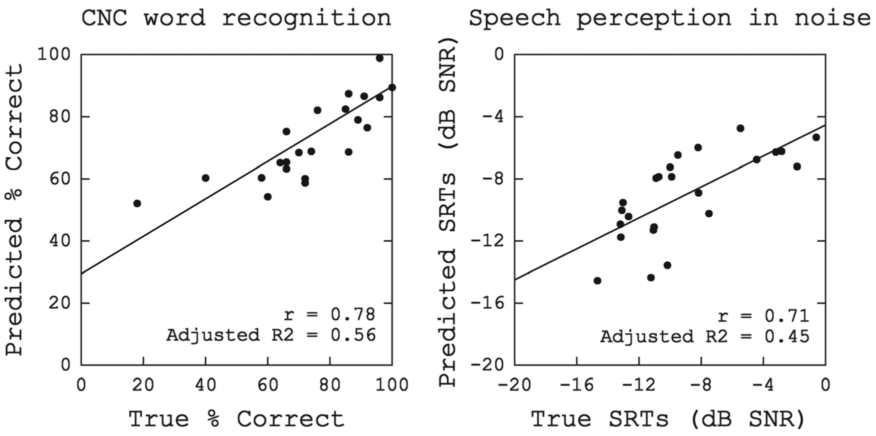

FIG. 8. Left panel shows the multiple regressions for CNC word recognition using mean modulation detection thresholds (MDTs) across the seven modulation frequencies and spectral-ripple thresholds ($r = 0.78$, adjusted $R^2 = 0.56$, $p < 0.0001$; whereas, independently, $R^2 = 0.42$ and $R^2 = 0.37$ for mean MDTs and spectral-ripple discrimination thresholds, respectively). Right panel shows the multiple regressions for speech perception in steady-state noise using mean MDTs and spectral-ripple thresholds ($r = 0.71$, adjusted $R^2 = 0.45$, $p < 0.0001$; whereas, independently, $R^2 = 0.32$ and $R^2 = 0.34$ for mean MDTs and spectral-ripple discrimination thresholds, respectively). Linear regressions between the true and predicted scores are represented by the solid lines.

The electrode outputs in Fig. 4 showed that the amount of modulation changed little when the AGC was turned off. The maximum current values also showed little change (Table II). The Advanced Bionics device uses a dual-level slow acting (325 ms attack at 57 dB sound pressure level) and fast acting (less than 0.6 ms attack at 65 dB sound pressure level) compression. The left panel of Fig. 3 shows that when the AGC was on, the MDTs at three different presentation levels (50, 65, and 75 dBA) for seven subjects were nearly equal across three different modulation frequencies, suggesting that AGC would have little effect on the amount of modulation for the moderate levels (50–75 dBA) used in the experiments. AGC, however, might have a measurable effect at levels greater than 75 dBA (Fig. 5 and Table II). Since AGC is designed to prevent peak clipping, it would be expected that at higher levels, the AGC would reduce peak clipping and improve modulation detection.

Evaluation of the effect of AGC on MDT is an example of a potential clinical use of the test. However, the modulation detection test with 7 frequencies for 6 tracking histories required too much testing time (2–2.5 hours). To make the test useful for clinical trials, for example, to determine the effect of mapping parameters or different sound processing strategies on MDT, it is important to use a test which is fast and efficient. For this purpose, the data were reanalyzed to evaluate the TMTFs with fewer tracking histories. MDTs for the first and second tracking history were averaged for each modulation frequency. If the difference of the two exceeded 3 dB, the final threshold for such condition was the mean of the first, second, and third tracking history. A similar procedure has been used to measure temporal modulation detection performance in children (e.g., Grose *et al.*, 1993; Hall and Grose, 1994). To further reduce the duration of testing, some modulation frequencies could be dropped from measurement. A *post hoc* analysis using the Tukey test showed that MDTs at 75 Hz were not significantly different from MDTs at 50 and 100 Hz. MDTs at 200 Hz were not also significantly different from its neighboring frequencies, 150 and 300 Hz. Measuring MDTs for 5 different modulation frequencies with two or three tracking histories would reduce the testing time significantly (to 30–40 min) and still measure accurate TMTFs. Mean MDTs across 5 frequencies estimated with two or three tracking histories were significantly correlated with mean MDTs across 7 frequencies estimated with 6 tracking histories ($r = 0.96$, $p < 0.0001$). Figure 9 shows comparison between the TMTFs determined both ways.

C. Effect of number of channels on MDTs

The present study showed that MDTs at a modulation frequency of 50 Hz were not dependent on the number of channels provided. As indicated in the rationale for Experiment 3, as the number of channels increased, the effect of increasing modulation information from multiple electrodes might have been offset by the effect of decreased filter bandwidth. Channel interaction could also affect MDTs. Chatterjee and Oba (2004) showed that across-electrode envelope interactions occur in the presence of various types of maskers and it decreases the modulation detection abilities in CI users. There are considerable differences in the

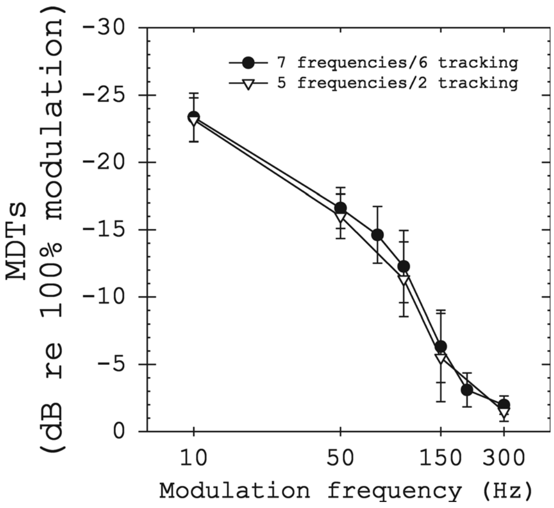

FIG. 9. Comparison of the TMTFs determined with 7 modulation frequencies and 6 tracking histories and 5 modulation frequencies and 2 or 3 tracking histories.

experimental conditions between the present study and the study by Chatterjee and Oba, which makes it difficult to interpret the present data in terms of across-electrode envelope interactions. For example, Chatterjee and Oba evaluated modulation detection performance in presence of maskers, which are not coherent to the signal. In contrast, as more channels were added in the present study, coherent modulation information was presented through multiple channels. The data showed that when the sound was presented through the sound processor, modulation detection sensitivity was neither increased nor decreased with more channels.

D. Relation between MDTs and speech perception abilities

The present study demonstrated that, when the CI listeners were tested with the sound processors both for speech recognition and modulation detection tasks, temporal modulation sensitivity was correlated with speech perception ability in CI users. The results are largely consistent with the previous reports, which demonstrated significant correlations between MDTs measured with single-electrode stimulation and speech perception measured with the sound processor (Cazals *et al.*, 1994; Fu 2002; Luo *et al.*, 2008). These previous studies with single-electrode, direct stimulation assessed subject's temporal sensitivity independent of the sound processor setting and still found significant relationship with speech perception ability. With clinical sound processors, the present study assessed listeners' temporal sensitivities and sound processor settings. When different sound encoding strategies were controlled, significant correlations between MDTs and speech perception scores were still found. Taken together, both from a single-electrode and multi-channel sound processing perspective, it is important to deliver accurate temporal modulation information for good speech processing in CI listeners.

The rate of decay of the TMTFs was also correlated with speech perception abilities ($r = 0.53$, $p = 0.008$ for CNC word scores; $r = -0.58$, $p = 0.003$ for SRTs in noise), indicating that individual subjects with steeper slopes of the TMTFs showed worse speech perception performance. This is consistent with a previous report by Cazals *et al.* (1994) where they demonstrated significant correlations between the slope of the function and vowel and consonant recognition. In general, the slopes of the TMTFs were primarily determined by the MDTs at higher modulation frequencies. For example, the correlations of the exponent b in the exponential function [Eq. (1)] with MDTs at 200 Hz and 300 Hz were significantly higher ($r = -0.79$ for 200 Hz and $r = -0.88$ for 300 Hz, $p < 0.0001$) than the correlations with the MDTs at 10 Hz and 50 Hz ($r = -0.17$ for 10 Hz and $r = -0.15$ for 50 Hz, $p = 0.5$). Note that a similar pattern is also observed in Table V that the correlations of speech recognition were greater with MDTs at 200 Hz and 300 Hz than with MDTs at 10 Hz and 50 Hz. Although the correlation analyses do not necessarily establish causation, this observation suggests that CI listeners who have better temporal processing at higher modulation frequencies tend to have better speech perception.

E. Relation between temporal modulation detection and spectral-ripple discrimination

MDTs were not found to be correlated with spectral-ripple thresholds at any of the 7 modulation frequencies, but both temporal modulation detection and spectral ripple discrimination significantly correlated with speech perception. We speculate that these two measures may represent two complementary and independent psychophysical abilities related to speech perception ability in CI listeners. This is further supported by the results of the linear regression analyses, showing that more than half of the variance in CNC word recognition scores was accounted for by the two measures (Fig. 8).

Alternatively, it is possible that CI users might use different places of excitation for temporal modulation detection and spectral-ripple discrimination, so performance on the two tasks was not correlated. A substantial variability has been observed in temporal modulation detection (Pfingst *et al.*, 2008) and in spectral-ripple discrimination (Anderson *et al.*, 2011) across different locations of cochlea. Future studies are needed to test hypotheses such as whether the temporal and spectral tasks share the same "optimal" place and whether the "optimal" place for each task can drive the performance outcomes with the sound processor.

Chatterjee and Yu (2010) investigated the relation between electrode discrimination and temporal modulation detection in CI listeners. They found that (1) the modulation detection at 100 Hz was significantly correlated with electrode discrimination when tested at low levels (20%–30% of the dynamic range) with bipolar stimulation; (2) weak or no correlations were found when tested with monopolar stimulation; and (3) no correlations were found when the two measures tested at a higher stimulation level (40% of the dynamic range). In the present study, all CI subjects used monopolar stimulation for modulation detection and spectral-ripple discrimination and acoustic stimuli were presented at moderate levels (65 dBA). The presentation level exceeds 40% of the dynamic range, thus the present results are consistent with Chatterjee and Yu. It is difficult to make a direct comparison between the present study and the study by Chatterjee and Yu, because spectral-ripple discrimination may involve integrating information from multiple locations along the cochlea, but electrode discrimination attempts to measure spectral sensitivity of local groups of neurons. Nevertheless, the two studies are in good agreement, showing no correlations between temporal (modulation detection) and spectral tasks (spectral-ripple and electrode discrimination) when tested with monopolar stimulation at moderate listening levels.

Creating a processing or rehabilitation strategy that improves temporal modulation sensitivity probably would not result in improvement in another dimension of hearing such as spectral resolution, due to the lack of relation between temporal modulation detection and spectral-ripple discrimination. However, the two measures have been shown to be important psychophysical abilities related to successful speech perception and increase the predictive power for speech recognition when they are both combined. Therefore, an important direction for future studies is to determine if improvement in clinical outcomes can be achieved if a new sound processor design improves temporal or spectral sensitivity without causing a decline in either.

ACKNOWLEDGMENTS

We appreciate the dedicated efforts of our subjects. This study was supported by NIH Grants No. R01-DC007525, No. R01-DC010148, No. P30-DC04661, No. F31-DC009755, No. T32-DC000033, No. L30-DC008490, Cochlear Corporation, and an educational fellowship from Advanced Bionics Corporation. Neither company played any role in data analysis or the composition of this paper. We thank three anonymous reviewers for numerous helpful comments on previous versions of this paper.

Anderson, E. S., Oxenham, A. J., Kreft, H., Nelson, P. B., and Nelson, D. A. (**2011**). "Comparing spectral tuning curves, spectral ripple resolution, and speech perception in cochlear implant users," J. Acoust. Soc. Am. (in press).

Bacon, S. P., and Viemeister, N. F. (**1985**). "Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners," Audiology **24**, 117–134.

Busby, P. A., Tong, Y. C., and Clark, G. M. (**1993**). "The perception of temporal modulations by cochlear implant patients," J. Acoust. Soc. Am. **94**, 124–131.

Carlyon, R. P., Deeks, J. M., and McKay, C. M. (**2010**). "The upper limit of temporal pitch for cochlear-implant listeners: Stimulus duration, conditioner pulses, and the number of electrodes stimulated," J. Acoust. Soc. Am. **127**, 1469–1478.

Cazals, Y., Pelizzone, M., Saudan, O., and Boex, 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 Oba, S. I. (**2004**). "Across- and within-channel envelope interactions in cochlear implant listeners," J. Assoc. Res. Otolaryngol. **5**, 360–375.

Chatterjee, M., and Yu, J. (**2010**). "A relation between electrode discrimination and amplitude modulation detection by cochlear implant listeners," J. Acoust. Soc. Am. **127**, 415–426.

Donaldson, G. S., and Viemeister, N. F. (**2000**). "Intensity discrimination and detection of amplitude modulation in electric hearing," J. Acoust. Soc. Am. **108**, 760–763.

Eddins, D. A. (**1993**). "Amplitude modulation detection of narrow-band noise: Effects of absolute bandwidth and frequency region," J. Acoust. Soc. Am. **93**, 470–479.

Firszt, J. B. (**2003**). *HiResolution^{TM} Sound Processing* (Advanced Bionics, LLC, Valencia, CA).

Fu, Q. J., Zeng, F. G., Shannon, R. V., and Soli, S. D. (**1998**). "Importance of tonal envelope cues in Chinese speech recognition," J. Acoust. Soc. Am. **104**, 505–510.

Fu, Q. J., and Shannon, R. V. (**2000**). "Effect of stimulation rate on phoneme recognition by nucleus-22 cochlear implant listeners," J. Acoust. Soc. Am. **107**, 589–597.

Fu, Q. J., and Zeng, F. G. (**2000**). "Identification of temporal envelope cues in Chineses tone recognition," Asia Pac. J. Speech Lang. Hear. **5**, 45–57.

Fu, Q. J. (**2002**). "Temporal processing and speech recognition in cochlear implant users," Neuroreport **13**, 1635–1639.

Fu, Q. J., Chinchilla, S., and Galvin, J. J. (**2004**). "The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users," J. Assoc. Res. Otolaryngol. **5**, 253–260.

Galvin, J. J., III, and Fu, Q. J. (**2005**). "Effects of stimulation rate, mode and level on modulation detection by cochlear implant users," J. Assoc. Res. Otolaryngol. **6**, 269–279.

Galvin, J. J., III, and Fu, Q. J. (**2009**). "Influence of stimulation rate and loudness growth on modulation detection and intensity discrimination in cochlear implant users," Hear. Res. **250**, 46–54.

Grose, J. H. Hall, J. W., and Gibbs, C. (**1993**). "Temporal analysis in children," J. Speech. Hear. Res. **36**, 351–356.

Hall, J. W., and Grose, J. H. (**1994**). "Development of temporal resolution in children as measured by the temporal modulation transfer functions," J. Acoust. Soc. Am. **96**, 150–154.

Harris, R. W. (**1991**). *Speech Audiometry Materials Compact Disk* (Brigham Young University, Provo, UT).

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

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

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

Litvak, L. M., Spahr, A. J., Saoji, A. A., and Fridman, G. Y. (**2007**). "Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners," J. Acoust. Soc. Am. **122**, 982–991.

Lomax, R. G. (**2005**). *Statistical Concepts: A Second Course* (Lawrence Erlbaum Associates, Mahwah, NJ).

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.

Luo, X., Galvin, J. J., and Fu, Q. J. (**2010**). "Effects of stimulus duration on amplitude modulation processing with cochlear implants," J. Acoust. Soc. Am. **127**, EL23–EL29.

McKay, C. M., Henshall, K. R., and Hull, A. E. (**2005**). "The effect of rate of stimulation on perception of spectral shape by cochlear implantees," J. Acoust. Soc. Am. **118**, 386–392.

McKay, C. M., and Henshall, K. R. (**2010**). "Amplitude modulation and loudness in cochlear implantees," J. Assoc. Res. Otolaryngol. **11**, 101–111.

Moore, B. C. J., Glasberg, B. R., Stone, M. A. (**1991**). "Optimization of a slow-acting automatic gain control system for us in hearing aids," Br. J. Audiol. **25**, 171–182.

Peterson, G. E., and Lehiste, I. (**1962**). "Revised CNC lists for auditory tests," J. Speech Hear. Disord. **27**, 62–70.

Pfingst, B. E., Xu, L., and Thompson, C. S. (**2007**). "Effects of carrier pulse rate and stimulation site on modulation detection by subjects with cochlear implants," J. Acoust. Soc. Am. **121**, 2236–2246.

Pfingst, B. E., Burkholder-Juhasz, R. A., Xu, L., and Thompson, C. S. (**2008**). "Across-site patterns of modulation detection in listeners with cochlear implants," J. Acoust. Soc. Am. **123**, 1054–1062.

Richardson, L. M., Busby, P. A., and Clark, G. M. (**1998**). "Modulation detection interference in cochlear implant subjects," J. Acoust. Soc. Am. **104**, 442–452.

Saoji, A. A., Litvak, L. M., Spahr, A. J., and Eddins, D. A. (**2009**). "Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners," J. Acoust. Soc. Am. **126**, 955–958.

Shannon, R. V. (**1992**). "Temporal modulation transfer functions in patients with cochlear implants," J. Acoust. Soc. Am. **91**, 2156–2164.

Turner, C. W., Gantz, B. J., Vidal, C., Behrens, A., and Henry, B. A. (**2004**). "Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing," J. Acoust. Soc. Am. **115**, 1729–1735.

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.

Viemeister, N. F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds," J. Acoust. Soc. Am. **66**, 1364–1380.

Wei, C., Cao, K., Jin, X., Chen, X., and Zeng, F. G. (**2007**). "Psychophysical performance and Mandarin tone recognition in noise by cochlear implant users," Ear Hear. **28**, 62S–65S.

Won, J. H., Drennan, W. R., and Rubinstein, J. T. (**2007**). "Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users," J. Assoc. Res. Otolaryngol. **8**, 384–392.

Won, J. H., Drennan, W. R., Kang, R. S., and Rubinstein, J. T. (**2010**). "Psychoacoustic abilities associated with music perception in cochlear implant users," Ear Hear. **31**, 796–805.

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.

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.