Brochier, T., H. J. McDermott, and C. M. McKay. 2017. The effect of presentation level and stimulation rate on speech perception and modulation detection for cochlear implant users. *The Journal of the Acoustical Society of America*. **141**(6): 4097.

The effect of presentation level and stimulation rate on speech perception and modulation detection for cochlear implant users

Running head:

Cochlear implant speech perception: rate effects

Tim Brochier, Hugh McDermott, Colette McKay

The Bionics Institute, 384-388 Albert Street, East Melbourne, Victoria 3002, Australia

University of Melbourne, Department of Medical Bionics, 384-388 Albert Street, East Melbourne, Victoria 3002, Australia

In order to improve speech understanding for cochlear implant (CI) users, it is important to maximize the transmission of temporal information. The combined effects of stimulation rate and presentation level on temporal information transfer and speech understanding remain unclear. The present study systematically varied presentation level (60, 50, and 40 dBA) and stimulation rate (500 and 2400 pulses per second per electrode (pps)) in order to observe how the effect of rate on speech understanding changes for different presentation levels. Speech recognition in quiet and noise, and acoustic amplitude modulation detection thresholds (AMDTs) were measured with acoustic stimuli presented to speech processors via direct audio input (DAI). With the 500 pps processor, results showed significantly better performance for CNC words in quiet, and a reduced effect of noise on sentence recognition. However, no rate or level effect was found for AMDTs, perhaps partly because of amplitude compression in the sound processor. AMDTs were found to be strongly correlated with the effect of noise on sentence perception at low levels. These results indicate that AMDTs, at least when measured with the CP910 Freedom speech processor via DAI, explain between-subject variance of speech understanding, but do not explain within-subject variance for different rates and levels.

I. INTRODUCTION

The perception of temporal amplitude modulations is critical for speech understanding. Cochlear implant users are especially reliant upon temporal information in speech, due to the limited spectral resolution of the cochlear implant. In order to transmit temporal information, cochlear implant processors separate the incoming acoustic signal into several frequency bands, each allocated to a specific electrode in the implanted array. In the most common signal processing strategies, the electrodes are activated with fixed-rate biphasic pulse trains, which are amplitude-modulated by the temporal envelope of their respective frequency band. The aim of the present study was to examine the effect of stimulation rate on speech perception for cochlear implant users, and to determine how the effect of rate on speech understanding changes for different presentation levels.

In the following, the stimulation rate is stated as the rate programmed in the sound processor, corresponding to the maximum rate on each active electrode, with units of pulses per second (pps).There is a wide range of available stimulation rates in current processors, from rates as low as 200 pps, to rates as high as 5000 pps. Assuming equal sampling rates of the microphone signal (typically around 16 kHz) and similar envelope extraction methods, high stimulation rates provide more detailed sampling of the temporal envelope than low stimulation rates. High rates also promote stochastic responses in auditory neurons (Rubinstein *et al.*, 1999), reducing unnatural phase locking observed in rates below 800 pps (Dynes and Delgutte, 1992). However, the presumed advantages of high stimulation rates may be offset by increased channel interaction (McKay *et al.*, 2005; Middlebrooks, 2004) or higher variation in perceived loudness cues (Azadpour *et al.*, 2015).

Many studies have examined the effect of stimulation rate on speech perception for cochlear implant users, with variable results. Using the CIS stimulation strategy, both Kiefer *et al.*

(2000) and Loizou *et al.* (2000) observed improved word and consonant recognition as pulse rates increased from 250 – 2000 pps. However, using the same strategy, Lawson *et al.* (1996) showed no effect of pulse rate for rates from 250 – 2525 pps, and Fu and Shannon (2000) only showed improved speech performance up to 150 pps, with insignificant differences for pulse rates from 150 - 500 pps.

Using the ACE stimulation strategy, Holden *et al.* (2002) tested rates between 720 and 1800 pps at presentation levels of 50, 60, and 70 dB SPL. While better group mean scores were measured for sentence recognition at 50 dB SPL using the high rate of 1800 pps compared to the low rate of 720 pps, all other presentation levels and speech tests (phonemes, words) showed no significant rate effects. Friesen *et al.* (2005) measured phoneme, word, and sentence recognition with the Clarion, Clarion II, and Nucleus 24, all using the CIS strategy. Rates from 200 – 5000 pps were compared and no significant differences were found between rates. Similarly, Weber *et al.* (2007) found that rates between 500 and 2500 pps had no influence on speech recognition scores for monosyllables or sentences, when using the ACE strategy on the Nucleus Freedom processor.

In a longitudinal study, Plant *et al.* (2007) tested the Nucleus 24 processor with the ACE strategy, and found that rate preference was highly subject dependent. Subjects chose a medium stimulation rate between 1200 and 1400 pps, and a high stimulation rate between 2400 and 3200 pps, and then completed speech perception tests and reported their preferences. Of 15 subjects, 5 preferred the medium rate, 8 preferred the high rate, and 2 were undecided. Only 2 subjects performed better in tests of both CNC words and CUNY sentences with the higher rate. Arora *et al.* (2009) tested the CI24 Contour implant with the ESPirit 3G Processor, and also found results were subject dependent. Group mean scores were significantly better for medium rates of 500 pps and 900 pps than for the low rate of 275 pps. Of 8 subjects, 3 were best at 500 pps, 3 at 900 pps, and 2 showed no significant difference between 500 and 900 pps. Shannon *et al.* (2011) came to similar conclusions, showing that for pulse rates from 600 – 5000 pps, CNC words and IEEE sentences in quiet and in noise had similar results across rates. They used the Advanced Bionics CII Processor with the CIS strategy. Park *et al.* (2012) showed significant rate effects with Korean sentences, with subjects performing better at the low-mid rate of 900 pps than the high rate of 2400 pps using the Nucleus 24 processor and the ACE strategy.

Nearly all of the above studies used presentation levels between 60 dB SPL and 70 dB SPL, which output currents in the upper half of the electrical dynamic range. At these high levels, there has been little rate effect shown in either speech, or in psychophysical correlates of speech, such as modulation detection thresholds and temporal modulation transfer functions. However, at lower levels (below 50% of the dynamic range), in studies using direct electrical stimulation, low rates consistently lead to better modulation detection thresholds than high rates (Fraser and McKay, 2012; Galvin III and Fu, 2005; Galvin and Fu, 2009; Green *et al.*, 2012; Pfingst *et al.*, 2007). Since modulation detection thresholds have been shown to correlate with speech perception ability (De Ruiter *et al.*, 2015; Fu, 2002; Gnansia *et al.*, 2014; Luo *et al.*, 2008; Won *et al.*, 2011), similar rate effects for speech at low levels could be hypothesized. That is, since modulation detection improves with low rates compared to high rates at low levels, it is of interest whether speech perception also improves for low rates compared to high rates at low levels.

Park (2012) and Holden (2002) are the only researchers to perform speech recognition tests at levels below 60 dB SPL, and have obtained conflicting results. Holden did not observe consistent differences in speech understanding between the high rate of 1800 pps and the low rate of 720 pps, but some subjects had better speech perception in noise at the higher rate for the 50 dB SPL stimulus. Park (2012), however, used presentation levels of 45 dB SPL, and found that subjects consistently performed better on Korean sentences and phonemes with the lower rate of 900 pps compared to the higher rate of 2400 pps.

The present study systematically varied presentation level and stimulation rate in order to observe how the effect of rate on speech understanding changes for different presentation levels. Word recognition in quiet, the effect of noise on sentence perception, and acoustic amplitude modulation detection thresholds (AMDTs) were measured. All measurements were performed using novel speech processor MAPs with rates of 500 pps and 2400 pps, at presentation levels of 40 dBA, 50 dBA, and 60 dBA. A two-way repeated measures analysis of variance (ANOVA) was used to test whether there was an interaction effect between stimulation rate and presentation level. It was hypothesized that speech understanding would be poor at the high rate compared to the low rate, only for stimuli at the lower level.

II. METHODS

A. Participants

Nine postlingually deafened adult CI users completed the study. Participants were recruited from the clinical population of the Royal Victorian Eye and Ear Hospital. Permission to conduct the studies was obtained from the Human Research and Ethics Committee of the Royal Victorian Eye and Ear Hospital, and each participant provided written informed consent. Participants were tested over the course of four to five sessions of 1-2 hours. Details about the participants are described in *Table I.*

B. Equipment

Participants were fit with the same CP910 Freedom Speech Processor. Two MAPs were created: Experimental MAP 1 with a rate of 500 pps, and Experimental MAP 2 with a rate of 2400 pps. The experimental MAPs used the ACE strategy with 6 maxima, pulse width of 25 μs, and interphase gap of 8.4 μs. The number of maxima was reduced from the clinical standard of 8 in order to keep the pulse width and interphase gap constant between the experimental MAPs (a pulse width of 25 μs with 8 maxima is unavailable for the rate of 2400 pps in the cochlear implant fitting software).

The standard clinical procedure at the Royal Victorian Eye and Ear Hospital was used to create the experimental MAPs. Threshold (T) and Comfort (C) levels were found for each electrode at both the 500 pps and 2400 pps rate. Loudness balancing was performed both at C levels and at 70% of the dynamic range, using three-electrode sweeps across the array. Participants were asked whether the loudness of the stimulation at each electrode was the same, and T and C levels were adjusted accordingly. More specifically, C-levels were adjusted during loudness balancing at C-level, and T-levels were adjusted during loudness balancing at 70% of the dynamic range. Finally, to ensure that both experimental MAPs were balanced in loudness, 50 dB SPL Bamford-Kowal-Bench (BKB) sentences (Bench *et al.*, 1979) were presented for both stimulation rates. C-levels were further adjusted to ensure that both experimental programs were loudness balanced. In order to have more control over the presentation level and processing of the stimuli, Adaptive Dynamic Range Optimization (ADRO), Autosensitivity Control (ASC), SmartSound iQ, Background Noise Reduction, Wind Noise Reduction, and Beamforming were all disabled.

The sensitivity was fixed at the default value of 12 for both experimental programs. The sensitivity control determines the minimum acoustic level in each channel that is mapped to the electrical T-level in that channel. The minimum acoustic level is both frequency-dependent and signal-dependent. At a sensitivity of 12, the minimum acoustic level for pure sine tones is 13 dB SPL for channels 1 through 8 (center frequencies 7438 down to 2875 Hz), and rises to 30 dB SPL at a slope of approximately 6 dB/octave for channels 9 through 22 (center frequencies 2300 down to 250 Hz). At the 40 dBA presentation level, some low level envelope cues were below the level that results in perceptible stimulation, which likely reduced the listeners' ability to understand speech in this condition. The sensitivity control also determines an upper acoustic threshold, above which all envelope levels are mapped to electrical C-level. This upper threshold is 40 dB above the minimum acoustic level in each channel. At the 60 dBA presentation level, the upper acoustic threshold would cause some high level envelope cues to be compressed, possibly affecting speech perception and modulation detection at this level.

The volume was fixed at 6 for the fitting and testing of both experimental MAPs. The volume control raises or lowers the electric C-levels by a certain percentage of the dynamic range. Fixing the volume at 6 for both fitting and testing ensured that the fitted C-levels for each experimental MAP remained unaltered through the duration of the experiment.

The participants' clinical MAP was not tested, because there would be a clear preference for the stimulation rate with which they are accustomed. The intent of this acute study was to observe immediate effects of altering the stimulation rate. The high rate of 2400 pps and low rate of 500 pps were chosen because they are above and below the usual rate of 900 pps for all participants, with the exception of P5, who uses a 250 pps clinical MAP. Thus both experimental MAPs were novel for all the participants.

C. Stimuli and Procedure

All speech and psychophysical stimuli were presented using the direct audio input of the CP910 Freedom processor in order to prevent the use of residual hearing. The presentation level to the direct audio input was calibrated to ensure that the output of the processor was equivalent for the same acoustic stimulus through the microphone input and direct audio input.

1. Speech Perception

Speech intelligibility was evaluated using Consonant-vowel Nucleus-Consonant (CNC) words in quiet and BKB sentences in quiet and in competing multi-talker babble noise. The CNC material comprised two lists, each containing 150 words recorded by a male Australian speaker. Each list comprised 50 different words at 3 presentation levels: 40 dBA (low), 50 dBA (mid-low), and 60 dBA (mid-high). The order of the words and levels in each list was randomized. For each test, the word list was selected at random, and no list was repeated for any subject. The participants were given one short practice list (16 words) with their clinical MAP before testing began. The participants were asked to repeat each word immediately after it was presented, and they were scored on correct phonemes identified out of a total of 150 phonemes at each level. They were tested with the two experimental MAPs (500 pps and 2400 pps).

The BKB sentences comprised one list of 640 sentences, each sentence containing three key words, recorded by an Australian male speaker. The participants were given one practice trial of 16 sentences before testing began. Aside from the practice lists, no other training was provided to participants for the novel MAPs. In order to minimize learning effects, the comparisons made in this study only were between equally novel MAPs, and the order of the testing was methodically varied among participants. Sentences were selected at random, and no sentence was repeated during any test for any subject. Participants were first presented with 32 sentences in quiet at each presentation level and each rate, in order to establish a baseline score in each condition. The participants were asked to repeat as many words as they could in each sentence. An adaptive procedure was used to assess the effect of competing noise on each listener’s ability to understand the sentences. The target SNR was the SNR at which the participant correctly identified 70% of the words that they recognized in the corresponding quiet conditions. SNR70% was measured (as in McKay and Henshall (2002) and McDermott *et al.* (2005)), as opposed to speech reception threshold (SNR for 50% correct), so that the measure reflected largely the effect of noise rather than being highly influenced by the performance in quiet. Initially the SNR was set to +15 dB. The speech was fixed at the level used in the quiet condition, and the noise level was adapted to change the SNR. In each trial, the subject was given three sentences, with a total of 9 key words. If they received a score higher than 70% of their quiet condition score, the SNR was lowered; conversely, if they received a score lower than 70% of their quiet condition score, the SNR was raised. The SNR was changed in steps of 5 dB until two reversals were obtained, then steps of 3 dB until six more reversals were obtained. The final SNR to achieve 70% of the words in quiet condition ($SNR_{70\%}$) was the average of these last 6 reversals. $SNR_{70\%}$ was evaluated for 3 presentation levels (40, 50, and 60 dBA) and 2 rates (500 pps and 2400 pps).

2. Acoustic Modulation Detection Threshold Estimation

AMDTs were measured at presentation levels of 60 dBA and 40 dBA using the 500 pps and the 2400 pps MAP. Stimuli were presented through the direct audio input of the CP910. Sinusoidal modulation was applied using *Equation 1*:

$$N_{Mod} = N_{Pink}\left(1 + M_{Depth} \cos\ (2\pi\ t\ f_{mod})\right) \quad (1)$$

N_{Mod} represents the modulated pink noise, and N_{Pink} represents the unmodulated pink noise. The M_{Depth} variable represents the modulation index, which was varied between 0.001 (essentially no modulation, -60 dB relative to 100% modulation) and 1 (full modulation, 0 dB relative to 100% modulation). The f_{mod} variable represents the modulation frequency and t represents time. The MAPs used for measuring AMDTs were the same as the 500 pps and 2400 pps experimental MAPs explained before, but with only the six lowest frequency channels activated (electrodes 17-22:). The CP910 requires at least 12 channels out of 22 to be enabled when programming a MAP, so the 6-channel MAPs were created by disabling channels 7-16, and setting the T and C levels of channels 1-6 to zero. This process resulted in an automatic reassignment of the frequency allocation in the processor. The bin widths for channels 17-21 were expanded from 125 Hz to 250 Hz, while the bin width for channel 22 remained 125 Hz. The range of center frequencies for channels 17-22 was expanded from 150-875 Hz to 150-1500 Hz. The pink noise stimuli were lowpass filtered with a 4th order Butterworth filter at 1500 Hz so that channels 1-6 were never selected as maxima in the ACE processing scheme. Recordings of the pulse parameters at the output of the speech processor verified that the stimulation rates for the 12-channel MAPs remained 500 pps and 2400 pps on electrodes 17-22. In this way, only channels 17-22 were activated by the noise stimuli. Using the 6-channel MAPs removed the effect of modulation on electrode selection, isolating modulation sensitivity as the factor influencing the AMDT measurement.

Once the 6-channel MAPs were created, loudness balancing was performed to ensure that the unmodulated stimuli were the same loudness as the modulated stimuli for each presentation level and stimulation rate. McKay and Henshall (2010) showed that modulated stimuli are perceived as louder by CI users than unmodulated stimuli with the same average current level. Therefore loudness cues can be used when identifying the modulated stimulus among unmodulated stimuli, rather than actual modulation detection. In the loudness balancing procedure, we presented a reference (unmodulated) and a test (modulated) stimulus to the participant, with controls to turn the level of the test stimulus up or down, in large steps of ±2 dB and small steps of ±0.5 dB.

Two trials of loudness balancing were performed at each of the presentation levels of 60 dBA and 40 dBA, stimulation rates of 500 pps and 2400 pps, and modulation depths of 0.05, 0.1, 0.2, and 0.4 (-26, -20, -14, and -8 dB relative to 100% modulation, respectively). The level of the test stimulus started at a random value 3-7 dB below the reference for trial 1, and 3-7 dB above the reference for trial 2. For trial 1, participants were asked to raise the level of the test stimulus in 2 dB steps until it was louder than the reference stimulus, and then to lower the level in 0.5 dB steps to make the stimuli match in loudness. For trial 2, participants were asked to lower the level of the test stimulus in 2 dB steps until it was quieter than the reference stimulus, and then to raise the level in 0.5 dB steps to make the stimuli match in loudness. The average of the final levels in the two trials was used to determine the level at which the stimuli were balanced. Interpolation was used in the adaptive AMDT procedure to determine the amount of stimulus level adjustment required as a function of modulation depth, similar to Galvin *et al.* (2014), to keep the modulated and unmodulated stimuli equal in loudness.

For the AMDT measurement, a three interval forced choice, adaptive two-down one-up procedure was used. Using this method, the modulation depth at which the participant correctly identified the modulated stimulus 71% of the time (Levitt, 1971) was found. Stimuli were all 500 ms bursts of modulated or unmodulated noise. The modulation frequency was 10 Hz, which is representative of temporal envelope cues in speech (Rosen, 1992). The 10 Hz modulated stimulus went through exactly 5 cycles between highest and lowest levels over the course of a 500 ms stimulus. During each trial, participants were presented with two unmodulated stimuli and one modulated stimulus in a randomized order, separated by 500 ms silence, with the task of identifying the modulated stimulus. If the subject correctly identified two modulated stimuli in a row, the modulation depth was reduced. If the subject incorrectly identified a modulated stimulus once, the modulation depth increased. For the first two reversals, a step size of ±6 dB re 100% modulation was used. For the next six reversals, a step size of ±2 dB re 100% modulation was used. The AMDT was determined by averaging the last six reversals.

In addition to applying the interpolated loudness balance to the modulated stimulus in each trial, level jitter in each interval was applied to remove the influence of any remaining loudness cues. The amount of jitter was determined for each subject using the variance in loudness balancing trials, and by the method explained in Dai and Micheyl (2010) when using a three-interval oddity forced choice task. The maximum 95% confidence interval of the standard error between the two loudness balancing trials across the different conditions was used to determine jitter range. Across all subjects and conditions, the minimum jitter range used was ±0.25 dB and the maximum jitter range used was ±2.2 dB.

III. RESULTS

Figure 1 shows CNC phoneme scores for different stimulation rates and presentation levels. A two-way repeated measures analysis of variance (ANOVA) was performed to assess the effect of stimulation rate and presentation level on speech perception. The ANOVA revealed a significant effect of level ($F(2,8)=59.86$, $p < 0.001$), with participants achieving better scores at louder levels. There was also a small but significant effect of rate ($F(1,8)=5.94$, $p = 0.019$), with speech understanding better at low rates. The rate effect was particularly evident at the lower level, where 7 out of 9 participants achieved a higher score with the low rate than with the high rate. However, no interaction effect between rate and level was found ($F(2,8)=1.43$, $p=0.251$). While the low rate advantage is statistically significant, it is not likely to be clinically significant. The maximal mean difference (across subjects) between CNC phoneme scores for the high rate and low rate occured at a presentation level of 40 dBA, and was only 6 percentage points. In addition, one participant (P6) used a very low clinical stimulation rate of 250 pps in his usual processor, as opposed to the middle stimulation rate of 900 pps for the other participants, and their results may have been better for 500 pps simply because it was closer to what they were accustomed to. When the results from this participant were removed, the repeated measures ANOVA revealed no significant difference between rates ($F(1,7)=3.15$, $p = 0.085$).

Table II shows the sentence in quiet scores for each participant in each condition. Figure 2 shows the $\mathrm{SNR}_{70\%}$ for different stimulation rates and presentation levels. The repeated measures ANOVA revealed a significant effect of rate for $\mathrm{SNR}_{70\%}$ (F(1,8)=11.56, p=0.002), with the lower rate of 500 pps consistently leading to better $\mathrm{SNR}_{70\%}$ than the higher rate of 2400 pps. The effect remained significant (F(1,7)=7.51, p = 0.007) when the participant who uses a 250 pps clinical pulse rate was removed. This rate effect became more pronounced as the level was lowered from 60 dBA to 40 dBA, with the mean difference between $\mathrm{SNR}_{70\%}$ for the different rates increasing from 0.8 dB to 2.5 dB. However, the ANOVA showed no interaction effect between rate and level (F(2,8)=0.99, p=0.380). No significant effect of level was found (F(2,8)=1.51, p=0.233), which was expected; sentences in quiet scores were worse at low levels compared to high levels, so the 70% target was lower for the low levels than for the high levels. Figure 3 shows AMDTs for different stimulation rates and presentation levels, with error bars representing ±1 standard error of the mean for each condition. Again, a two-way repeated measures ANOVA was used to assess the effect of stimulation rate and presentation level on AMDT. For this task, no significant effect of rate (F(1,8)=0.12, p=0.729), level (F(1,8)=0.68, p=0.417), or interaction between rate and level (F(1,8)=1.16, p=0.293) was found.

Figure 4 shows the Pearson correlation analyses of AMDTs with CNC phoneme scores . Only one significant correlation was found between AMDT and CNC phoneme score, at a presentation level of 60 dBA and a pulse rate of 2400 pps (R = -0.761, p = 0.018). However, this correlation is mainly driven by an outlier who found the modulation detection task particularly difficult. When this outlier is removed, the correlation becomes insignificant for all conditions.

Figure 5 shows the Pearson correlation analyses of AMDTs with $\mathrm{SNR}_{70\%}$. Significant correlations were found between AMDT and $\mathrm{SNR}_{70\%}$ for all conditions. At the high levels, the correlations were mainly driven by the same outlier mentioned above. When the outlier was removed, the correlations at the high levels became insignificant. However, at the low level, correlations remained significant for both the 500 pps and 2400 pps stimulation rate (R=0.86 and P=0.006 for both rates), indicating a strong relationship between $SNR_{70\%}$ and AMDTs.

IV. DISCUSSION

The results suggest that there was some advantage for low rates compared to high rates, particularly in noisy conditions. There was a consistent and significant effect of rate for $SNR_{70\%}$, in favor of the 500 pps rate. While there was a trend for the low-rate advantage to increase at low levels, as we hypothesized, no significant interaction between rate and level was found for any stimulus.

However, the low rate advantage of $SNR_{70\%}$ cannot be attributed to better detection of modulations in the acoustic stimulus at low rates compared to high rates. Despite the correlation between AMDTs and $SNR_{70\%}$ scores, there was no significant difference between AMDTs at high and low rates at either level. The lack of degradation in AMDTs with stimulus level (in contrast to MDTs measured with direct stimulation) was consistent with Won *et al.* (2011), who found no significant difference between AMDTs at 75 dBA, 65 dBA, and 50 dBA using acoustic stimulation through the processor. In contrast, when using direct electrical stimulation for MDT measurement, lower levels have consistently led to worse MDTs, with the low rate generally outperforming the high rate at low levels (Chatterjee and Oba, 2005; Fraser and McKay, 2012; Fu, 2002; Galvin III and Fu, 2005; Galvin and Fu, 2009; Pfingst *et al.*, 2007; Pfingst *et al.*, 2008; Zhou and Pfingst, 2014).

One reason for AMDTs not getting worse at low levels was that there were two stages of compression in the CI signal processor that potentially influenced temporal modulations at high input levels. Each of these compression stages was analyzed to assess their influence on the results. In the Freedom speech processor, the first stage of compression was an automatic gain control (AGC) which operated across all channels. However, it was found that the AGC was not active for any of the modulated noise stimuli. A test stimulus was created by adding a constant low-level 7063 Hz sinusoid (center frequency of channel 1) to a modulated noise stimulus with a depth of -6 dB re 100% modulation. The 7063 Hz probe sine tone caused an approximately constant level pulse train at channel 1 of the processor output when the AGC was not active. When the AGC was active, the gain control reduced the level of the probe sine tone during peaks in the modulation cycle. The upper panel of figure 6 shows the output of channel 19, which was activated by the modulated noise stimulus (at a modulation depth of -6 dB re 100% modulation), for different presentation levels. The lower panel of the same figure shows the output of channel 1, which was activated by the 7063 Hz probe sine tone. Even at a modulation depth of -6 dB re 100% modulation, which was well above the modulation detection threshold for 8 of the 9 participants, the AGC only began to activate when the noise was centered at 68 dBA. Therefore, for our high level stimuli centered at 60 dBA, the AGC did not influence the results.

The next compression stage was the nonlinear mapping from acoustic level to electrical level in each individual channel in the speech processor. The mapping was influential at high levels, where all acoustic envelope levels above a certain upper threshold were mapped to C-level. In order to measure the mapping from acoustic level to electrical level, a 625 Hz sine tone (center frequency of channel 19) was delivered through the direct audio input at a range of presentation levels from 20 - 75 dB SPL in steps of 2 dB. For each 625 Hz sine acoustic input level, the electrical output of electrode 19 was recorded (figure 7).

The nonlinearity in the mapping meant that for the same modulation depth of an acoustic modulated noise stimulus, the resulting electrical stimulation had different electrical modulation depths at different levels. In order to quantify this effect, pulse trains at a modulation depth of 0.2 (-14 dB re 100% modulation) were measured from the output of the processor at presentation levels of 60 dBA and 40 dBA. This modulation depth was chosen because it was around the average AMDT measured across conditions. The processor was programmed with a test MAP with a stimulation rate of 500 pps and with all T-levels set to 150 CL steps and C-levels set to 200 CL steps. The average peak to valley difference in current level steps was approximated by taking the difference between the 95th percentile and the 5th percentile of current level step values during the stimulus. Values were expressed as percentage of the dynamic range (%DR) in *Table III*. For acoustic stimuli of equal modulation depth, the electrical modulation depths were greater at 40 dBA than at 60 dBA. The increase in modulation depth at the output of the processor for low compared to high input levels may have been offset by the decreased modulation sensitivity at low levels, leading to no effect of level for equal modulation depth at the acoustic input. The compression associated with CI signal processing explained the lack of level effect for AMDTs, but did not explain the lack of effect of rate on AMDTs. One reason no effect of rate was found may be that at higher rates, the electrical dynamic range was larger. Consequently, an acoustic input with a particular modulation depth was mapped to a larger electrical modulation depth for the high rate of 2400 pps than for the low rate of 500 pps. This effect would counteract the increased modulation sensitivity at low rates compared to high rates reported in direct electrical MDTs (Galvin III and Fu 2005, Pfingst, Xu et al. 2007, Galvin and Fu 2009, Fraser and McKay 2012, Green, Faulkner et al. 2012)). Fraser and McKay (2012) expressed direct electrical MDTs as a proportion of the dynamic range, showing that the effect of stimulation rate was reduced, and that the higher rate of 2400 pps only led to significantly poorer MDTs at the low level of 40% of the dynamic range and high modulation frequency of 150 Hz.

Another key difference between AMDTs and direct electrical MDTs is the number of active channels, which could also explain the lack of effect of rate on AMDTs measured through the CI processor. Usually, measurements of direct electrical MDTs have only activated one electrode at a time, while measurements of AMDTs in the present study activated six adjacent channels at a time. The adjacent channels stimulated partly overlapping nerve fibre populations. Since the electrodes were activated in an interleaved fashion, the overall pulse rate on the overlapping nerve fibre population was higher than the low rate of 500 pps, and even higher for the high rate of 2400 pps. In studies that compared direct electrical MDTs with different stimulation rates, significant differences were only observed between low rates (<800 pps) and high rates (>1000 pps), while there was generally no significant difference in MDTs between high rates above 1000 pps (Galvin and Fu, 2009; Green *et al.*, 2012). The increased number of active channels for AMDT measurement through the speech processor could have reduced the advantage for low rates compared to high rates using direct electrical, single-channel MDT measurement, because the stimulation of overlapping nerve fibre populations by six adjacent channels effectively raised the stimulation rates being compared in this study by up to six-fold.

Another noteworthy finding from the present study was the strong correlation between low level AMDTs and $\mathrm{SNR}_{70\%}$. A reason for the strong correlation may have been the similarity between identifying modulations in noise and perceiving speech in noise. When unmodulated noise was passed through the speech processor, modulations inherent in the noise were encoded. These inherent modulations made the modulation detection task more difficult.. In order to identify modulation in a noise stimulus, the listener needed to be able to distinguish a specific sinusoidal modulation among many random modulations on each active electrode. Similarly, in order to perceive speech in noise, the listener needed to be able to differentiate modulations in the target speech from modulations in the background noise. These temporal abilities could have been particularly important at low levels, where the spectral representation of speech was worse because low-amplitude speech cues were not encoded.

V. CONCLUSION

The effect of noise on speech understanding was greater at high rates compared to low rates. Despite the correlation between AMDTs and $\mathrm{SNR}_{70\%}$, better $\mathrm{SNR}_{70\%}$ with the lower rate cannot be attributed to better detection of modulations in the acoustic stimulus at the lower rate. This result would indicate that while detection of modulations in the acoustic stimulus can be used to explain differences between subjects, it cannot be used to explain differences within the same subject for different rate and level conditions. If the correlation between AMDTs and speech perception represents a causal relationship, speech processing strategies aimed toward improving the detection of temporal modulations could improve speech perception for cochlear implant users. However, this study suggests that the adjustment of stimulation rate, regardless of level, does not improve the detection of modulations in the acoustic stimulus presented through the speech processor.

ACKNOWLEDGEMENTS

We are grateful to our dedicated participants, who graciously volunteered time and energy in many test sessions. TB is sponsored by the Melbourne International Research Scholarship and the Melbourne International Fee Remission Scholarship. The research was supported by a Veski Fellowship to CMM. The Bionics Institute acknowledges support it receives from the Victorian Government through its Operational Infrastructure Support Program.

REFERENCES

Arora, K., Dawson, P., Dowell, R., and Vandali, A. (**2009**). "Electrical stimulation rate effects on speech perception in cochlear implants," International journal of audiology **48**, 561-567.

Azadpour, M., Svirsky, M., and McKay, C. (**2015**). "Why is Current Level Discrimination Worse at High Stimulation Rates?," Conference on Implantable Auditory Prostheses 2015

Bench, J., Kowal, Å., and Bamford, J. (**1979**). "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," British journal of audiology **13**, 108-112.

Chatterjee, M., and Oba, S. I. (**2005**). "Noise improves modulation detection by cochlear implant listeners at moderate carrier levelsa)," The Journal of the Acoustical Society of America **118**, 993-1002.

Dai, H., and Micheyl, C. (**2010**). "On the choice of adequate randomization ranges for limiting the use of unwanted cues in same-different, dual-pair, and oddity tasks," Attention, Perception, & Psychophysics **72**, 538-547.

De Ruiter, A. M., Debruyne, J. A., Chenault, M. N., Francart, T., and Brokx, J. P. (**2015**). "Amplitude Modulation Detection and Speech Recognition in Late-Implanted Prelingually and Postlingually Deafened Cochlear Implant Users," Ear and hearing **36**, 557-566.

Dynes, S. B., and Delgutte, B. (**1992**). "Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea," Hearing research **58**, 79-90.

Fraser, M., and McKay, C. M. (**2012**). "Temporal modulation transfer functions in cochlear implantees using a method that limits overall loudness cues," Hearing research **283**, 59-69.

Friesen, L. M., Shannon, R. V., and Cruz, R. J. (**2005**). "Effects of stimulation rate on speech recognition with cochlear implants," Audiology and Neurotology **10**, 169-184.

Fu, Q.-J., and Shannon, R. V. (**2000**). "Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners," The Journal of the Acoustical Society of America **107**, 589-597.

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

Galvin III, J. J., and Fu, Q.-J. (**2005**). "Effects of stimulation rate, mode and level on modulation detection by cochlear implant users," Journal of the Association for Research in Otolaryngology **6**, 269-279.

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

Galvin, J. J., Fu, Q.-J., Oba, S., and Başkent, D. (**2014**). "A method to dynamically control unwanted loudness cues when measuring amplitude modulation detection in cochlear implant users," Journal of neuroscience methods **222**, 207-212.

Gnansia, D., Lazard, D. S., Léger, A. C., Fugain, C., Lancelin, D., Meyer, B., and Lorenzi, C. (**2014**). "Role of slow temporal modulations in speech identification for cochlear implant users," International journal of audiology **53**, 48-54.

Green, T., Faulkner, A., and Rosen, S. (**2012**). "Variations in carrier pulse rate and the perception of amplitude modulation in cochlear implant users," Ear and hearing **33**, 221-230.

Holden, L. K., Skinner, M. W., Holden, T. A., and Demorest, M. E. (**2002**). "Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy," Ear and hearing **23**, 463-476.

Kiefer, J., von Ilberg, C., Hubner-Egner, J., Rupprecht, V., and Knecht, R. (**2000**). "Optimized speech understanding with the continuous interleaved sampling speech coding strategy in patients with cochlear implants: effect of variations in stimulation rate and number of channels," Annals of Otology, Rhinology & Laryngology **109**, 1009-1020.

Lawson, D. T., Wilson, B. S., Zerbi, M., and Finley, C. C. (**1996**). "Speech processors for auditory prostheses," Third quarterly progress report, NIH contract

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," The Journal of the Acoustical society of America **49**, 467-477.

Loizou, P. C., Poroy, O., and Dorman, M. (**2000**). "The effect of parametric variations of cochlear implant processors on speech understanding," The Journal of the Acoustical Society of America **108**, 790-802.

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 and hearing **29**, 957.

McKay, C. M., and Henshall, K. R. (**2002**). "Frequency-to-electrode allocation and speech perception with cochlear implants," The Journal of the Acoustical Society of America **111**, 1036-1044.

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," The Journal of the Acoustical Society of America **118**, 386-392.

McKay, C. M., and Henshall, K. R. (**2010**). "Amplitude modulation and loudness in cochlear implantees," Journal of the Association for Research in Otolaryngology **11**, 101-111.

Middlebrooks, J. C. (**2004**). "Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds," The Journal of the Acoustical Society of America **116**, 452-468.

Park, S. H., Kim, E., Lee, H.-J., and Kim, H.-J. (**2012**). "Effects of Electrical Stimulation Rate on Speech Recognition in Cochlear Implant Users," Korean Journal of Audiology **16**, 6-9.

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 implantsa)," The Journal of the Acoustical Society of America **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 implantsa)," The Journal of the Acoustical Society of America **123**, 1054-1062.

Plant, K., Holden, L., Skinner, M., Arcaroli, J., Whitford, L., Law, M.-A., and Nel, E. (**2007**). "Clinical evaluation of higher stimulation rates in the nucleus research platform 8 system," Ear and hearing **28**, 381-393.

Rubinstein, J., Wilson, B., Finley, C., and Abbas, P. (**1999**). "Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation," Hearing research **127**, 108-118.

Shannon, R. V., Cruz, R. J., and Galvin, J. J. (**2011**). "Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise," Audiology and Neurotology **16**, 113-123.

Weber, B. P., Lai, W. K., Dillier, N., von Wallenberg, E. L., Killian, M. J., Pesch, J., Battmer, R. D., and Lenarz, T. (**2007**). "Performance and preference for ACE stimulation rates obtained with nucleus RP 8 and freedom system," Ear and hearing **28**, 46S-48S.

Won, J. H., Drennan, W. R., Nie, K., Jameyson, E. M., and Rubinstein, J. T. (**2011**). "Acoustic temporal modulation detection and speech perception in cochlear implant listenersa)," The Journal of the Acoustical Society of America **130**, 376-388.

Zhou, N., and Pfingst, B. E. (**2014**). "Effects of site-specific level adjustments on speech recognition with cochlear implants," Ear and hearing **35**, 30-40.

TABLE I. Relevant information about the cochlear implant users who participated in the study.

	Gender	Age (years)	Duration of hearing loss before implantion (years)	Duration of implant use (years)	Etiology	Usual Stimulation Rate (pps/electrode)
P1	Male	44	5	5	Unknown, Genetic	900
P2	Male	57	23	7	Unknown, Genetic	900
P3	Male	70	13	7	Unknown, Genetic	900
P4	Female	64	10	6	Unknown	900
P5	Male	78	23	15	Genetic	250
P6	Male	73	14	1.5	Chronic ear infections	900
P7	Female	65	16	5	Unknown, progressive hearing loss	900
P8	Male	73	1	5	Partially due to noise exposure, the rest unknown	900
P9	Female	66	20	12	Genetic	900

TABLE II. Sentences in quiet score for each participant and each condition. Scores are the number of words correct out of 100.

	60 dBA, 500 pps	60 dBA, 2400 pps	50 dBA, 500 pps	50 dBA, 2400 pps	40 dBA, 500 pps	40 dBA, 2400 pps
P1	100	100	100	100	100	98
P2	100	99	100	99	97	98
P3	91	89	98	93	97	74
P4	80	78	81	79	63	76
P5	98	94	99	91	78	81
P6	100	98	98	96	95	91
P7	100	99	100	97	100	99
P8	84	86	93	87	88	67
P9	46	53	48	36	64	47

TABLE III. Average peak to valley current level step differences, measured as the difference between the 95th percentile and 5th percentile of current level steps in the modulated noise pulse train. Values are expressed as a percentage of the dynamic range, and in current level steps for a test MAP with all T-levels set to 150 CL steps and all C-levels set to 200 CL steps.

Stimulus	Electrode 17	Electrode 19	Electrode 21
60 dBA, mod depth 0.2	24.0% DR 12.0 CL steps	28.0% DR 14.0 CL steps	35.2% DR 17.6 CL steps
40 dBA, mod depth 0.2	37.2% DR 18.6 CL steps	35.6% DR 17.8 CL steps	38.0% DR 19.0 CL steps

FIG. 1. Mean CNC phoneme scores compared to stimulation rate and presentation level for 9 participants. The error bars represent ±1 standard error of the mean.

FIG. 2. Mean Signal to noise ratio (SNR) required to achieve 70% of the sentences in quiet score, versus stimulation rate and presentation level for 9 participants. The error bars represent ±1 standard error of the mean.

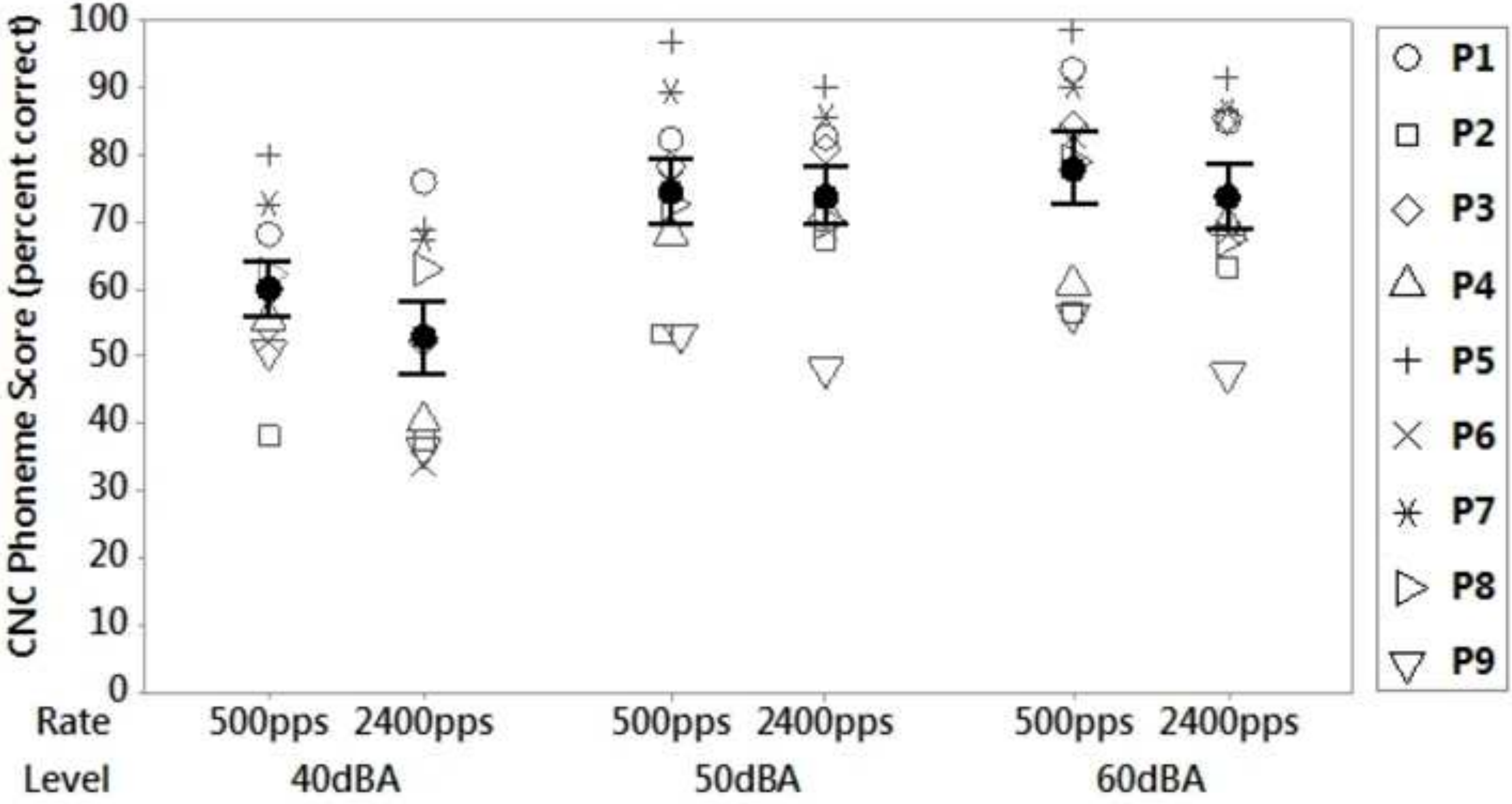
FIG. 3. Mean acoustic modulation detection thresholds, versus stimulation rate and presentation level for 9 participants. The error bars represent ±1 standard error of the mean.

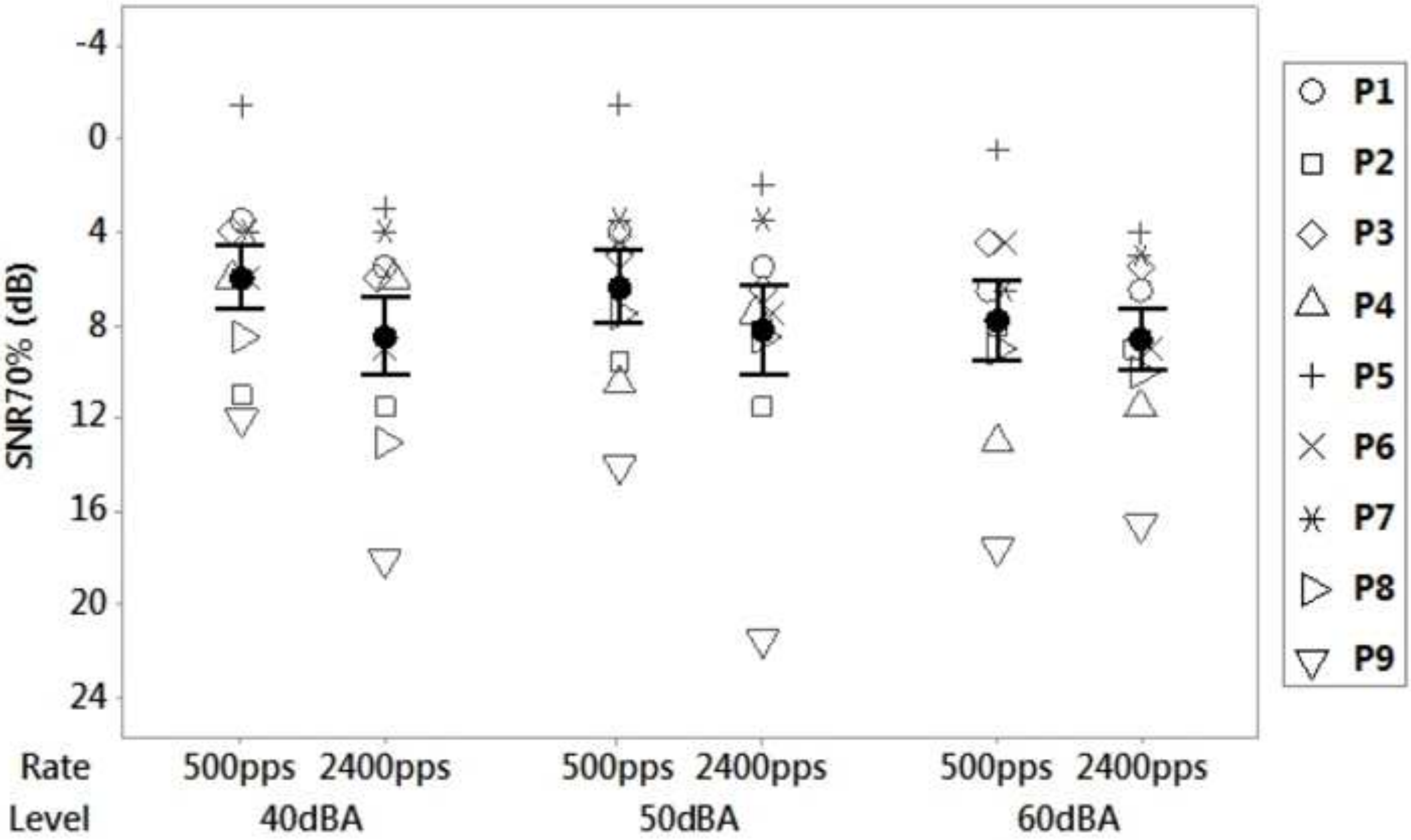
FIG. 4. Pearson correlation analyses between CNC phoneme scores and AMDTs for 9 participants, at different stimulation rates and presentation levels.

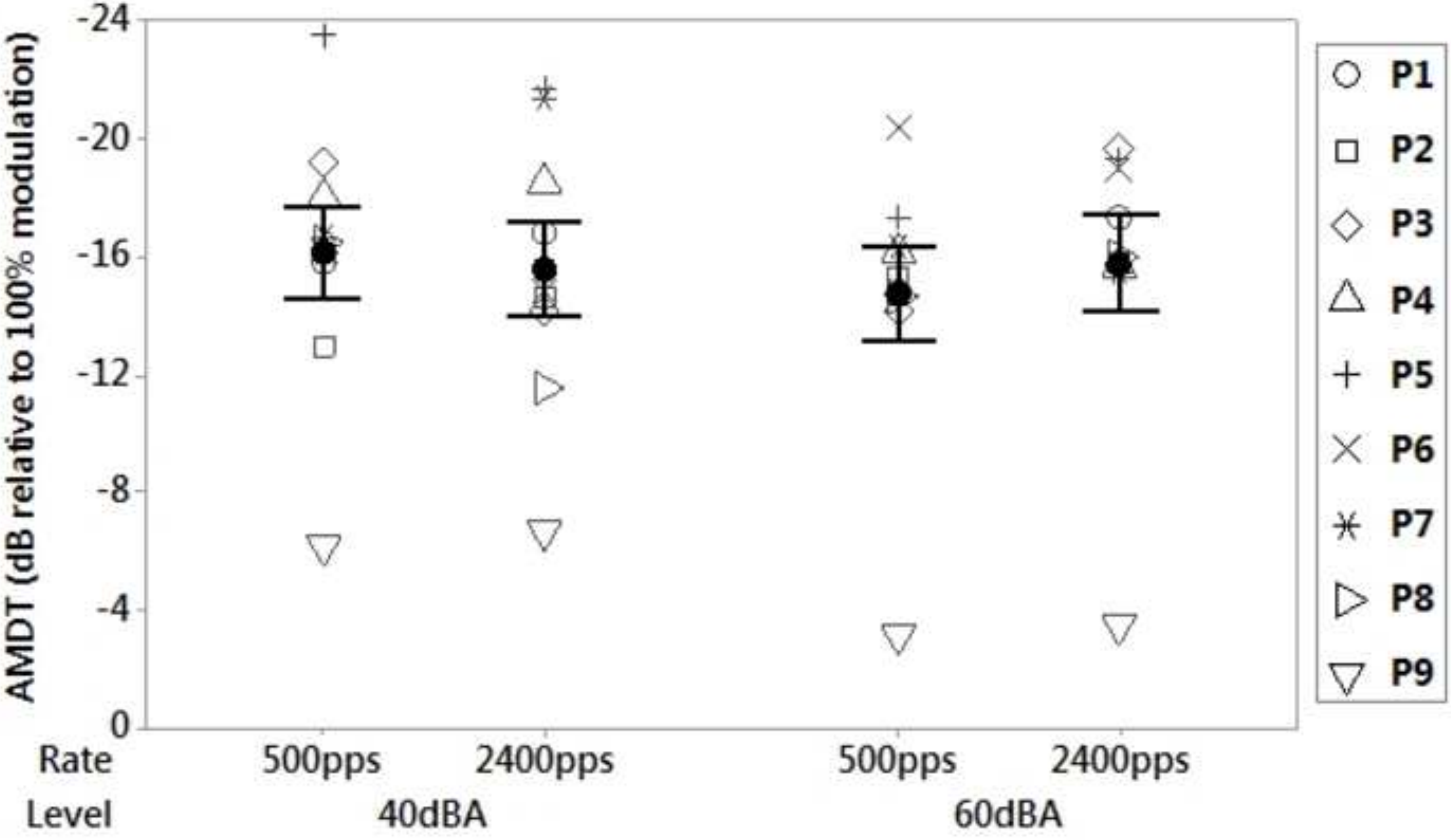
FIG. 5. Pearson correlation analyses between $\mathrm{SNR}_{70\%}$ and AMDTs for 9 participants, at different stimulation rates and presentation levels.

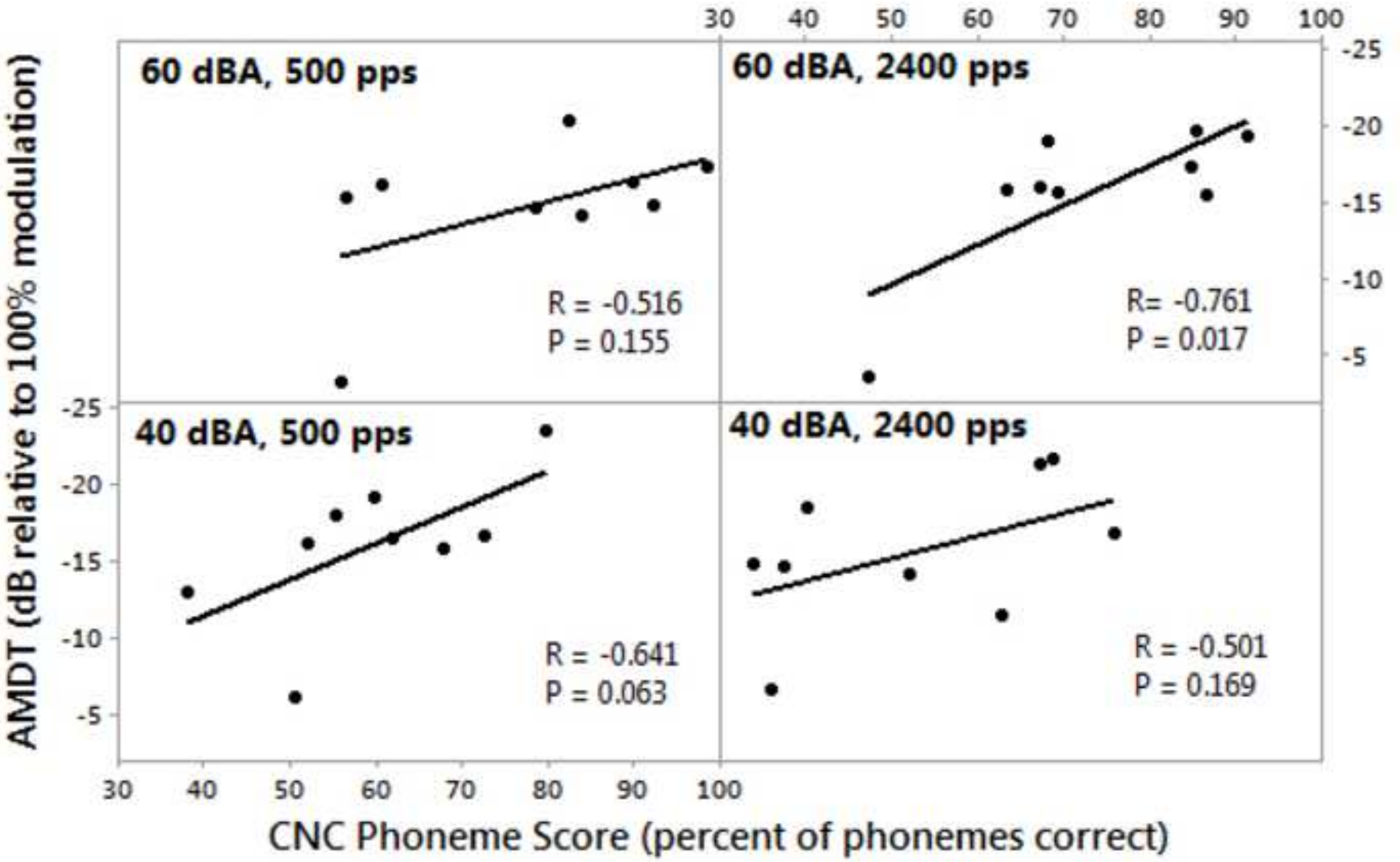
FIG. 6. Speech processor outputs at channel 19 (625 Hz center frequency, upper panel) and channel 1 (7063 Hz center frequency, lower panel), for a 500 ms test signal of 10 Hz modulated noise, lowpass filtered at 1500 Hz, plus a constant 7063 Hz sine tone. The modulation depth of the noise was -6 dB re 100% modulation.

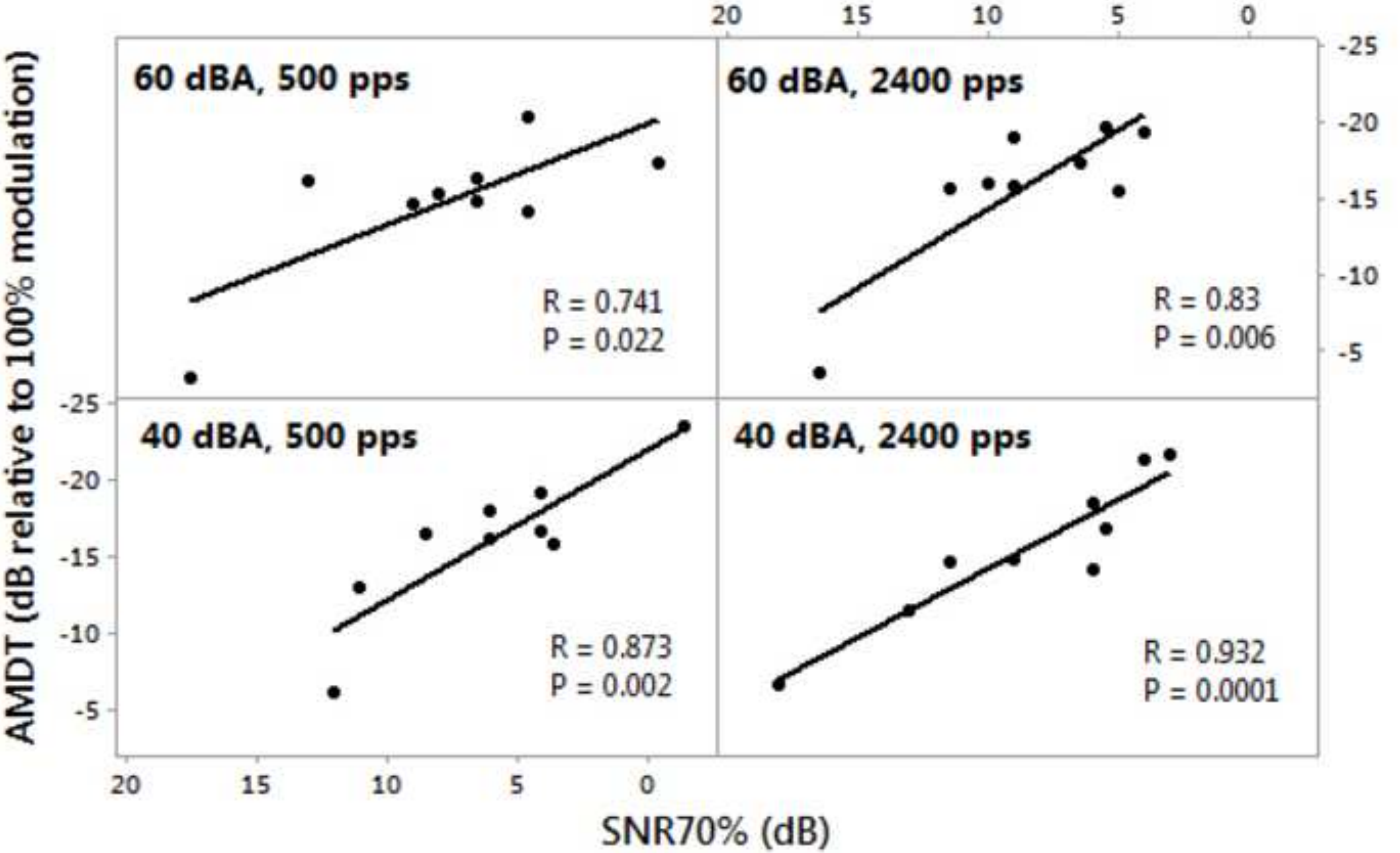
FIG. 7. Electrical output level at electrode 19 (625 Hz center frequency) for different acoustic input levels of a 625 Hz sine tone.


CNC Phoneme Score (percent correct)
100
90
80
70
60
50
40
30
20
10
0
Rate
500pps
2400pps
500pps
2400pps
500pps
2400pps
Level
40dBA
50dBA
60dBA
○ P1
□ P2
◇ P3
△ P4
+ P5
× P6
⚹ P7
▷ P8
▽ P9

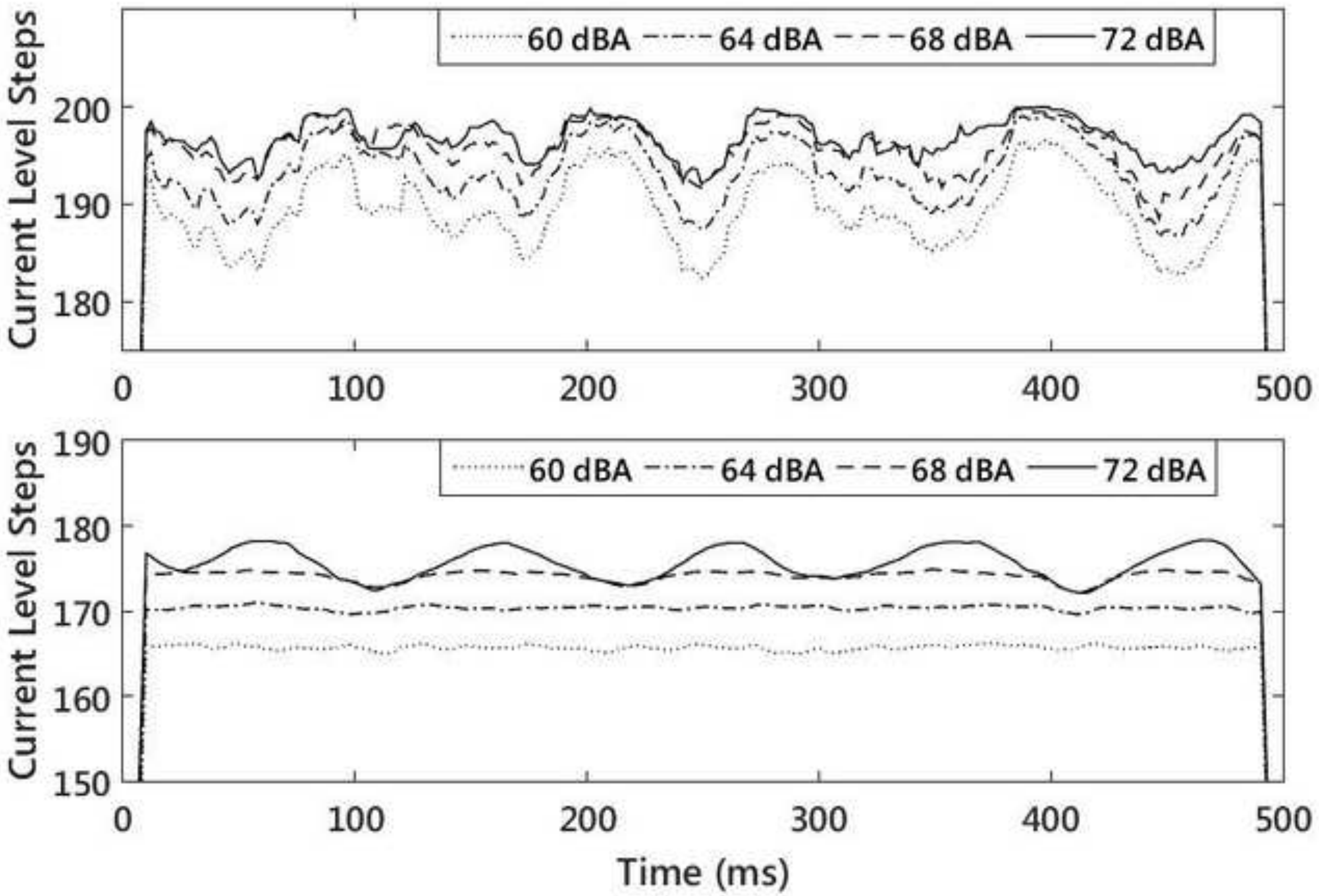

SNR70% (dB)
-4
0
4
8
12
16
20
24
Rate
500pps
2400pps
500pps
2400pps
500pps
2400pps
Level
40dBA
50dBA
60dBA
○ P1
□ P2
◇ P3
△ P4
+ P5
× P6
⋇ P7
▷ P8
▽ P9

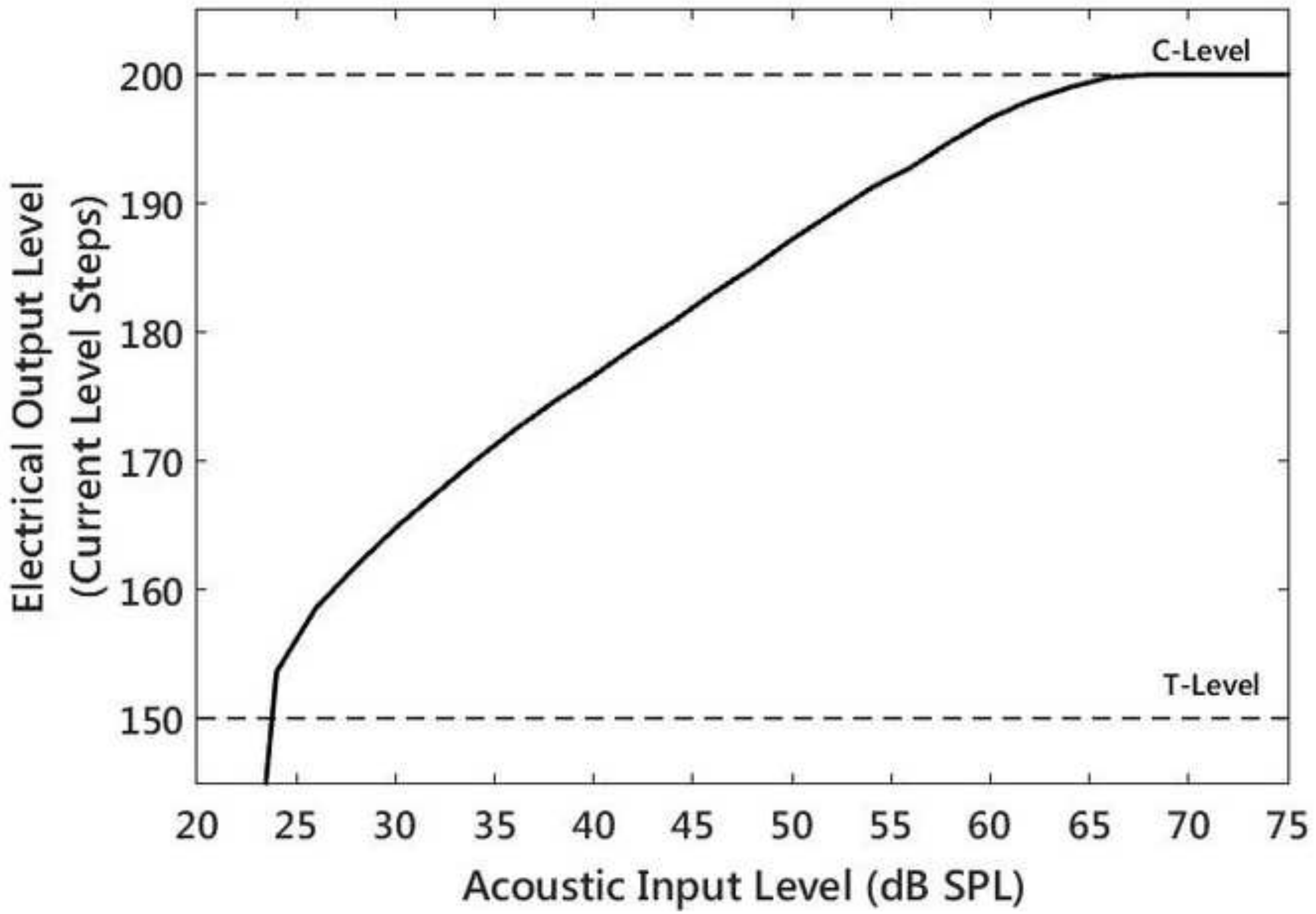

AMDT (dB relative to 100% modulation)
-24
-20
-16
-12
-8
-4
0
Rate
500pps
2400pps
500pps
2400pps
Level
40dBA
60dBA
○ P1
□ P2
◇ P3
△ P4
+ P5
× P6
⚹ P7
▷ P8
▽ P9


60 dBA, 500 pps
R = -0.516
P = 0.155
60 dBA, 2400 pps
R= -0.761
P = 0.017
40 dBA, 500 pps
R = -0.641
P = 0.063
40 dBA, 2400 pps
R = -0.501
P = 0.169
AMDT (dB relative to 100% modulation)
CNC Phoneme Score (percent of phonemes correct)
30
40
50
60
70
80
90
100
-25
-20
-15
-10
-5



60 dBA, 500 pps
R = 0.741
P = 0.022
60 dBA, 2400 pps
R = 0.83
P = 0.006
40 dBA, 500 pps
R = 0.873
P = 0.002
40 dBA, 2400 pps
R = 0.932
P = 0.0001
20
15
10
5
0
-25
-20
-15
-10
-5
AMDT (dB relative to 100% modulation)
SNR70% (dB)


60 dBA
64 dBA
68 dBA
72 dBA
Current Level Steps
200
190
180
0
100
200
300
400
500
60 dBA
64 dBA
68 dBA
72 dBA
Current Level Steps
190
180
170
160
150
0
100
200
300
400
500
Time (ms)


C-Level
T-Level
200
190
180
170
160
150
20
25
30
35
40
45
50
55
60
65
70
75
Electrical Output Level (Current Level Steps)
Acoustic Input Level (dB SPL)