Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users (L)

Heather A. Kreft,[a] Andrew J. Oxenham,[b] and David A. Nelson
Department of Otolaryngology, Clinical Psychoacoustics Laboratory, University of Minnesota, Rm 8-323
Phillips-Wangensteen Building, 420 Delaware Street Southeast, Minneapolis, Minnesota 55455

Detection and modulation rate discrimination were measured in cochlear-implant users for pulse-trains that were either sinusoidally amplitude modulated or were modulated with half-wave rectified sinusoids, which in acoustic hearing have been used to simulate the response to low-frequency temporal fine structure. In contrast to comparable results from acoustic hearing, modulation rate discrimination was not statistically different for the two stimulus types. The results suggest that, in contrast to binaural perception, pitch perception in cochlear-implant users does not benefit from using stimuli designed to more closely simulate the cochlear response to low-frequency pure tones. © *2010 Acoustical Society of America.* [DOI: 10.1121/1.3282947]

I. INTRODUCTION

Cochlear-implant (CI) users often have difficulty with speech recognition in noisy situations (e.g., Friesen *et al.*, 2001; Nelson *et al.*, 2003) as well as with pitch and music perception (e.g., McDermott, 2004; Gfeller *et al.*, 2007; Galvin *et al.*, 2007). A number of studies have suggested that pitch coding deficits might underlie some of the speech-perception difficulties faced by CI users in noisy situations by impairing CI users' ability to make use of F0-based segregation cues (Stickney *et al.*, 2004; Turner *et al.*, 2004; Qin and Oxenham, 2005). Some efforts have been made to improve F0 representation in CIs, so far with relatively little success (Geurts and Wouters, 2001; Green *et al.*, 2005; Vandali *et al.*, 2005; Laneau *et al.*, 2006). Most recently, Landsberger (2008) compared pitch discrimination using a number of different waveforms, including sine, sawtooth, sharpened sawtooth, and square waves. The results suggested no great benefit of any of the waveforms, leading Landsberger to conclude that the waveform shapes may be used interchangeably and their use in future speech processing strategies may ultimately depend on other engineering issues.

It remains possible that a more "optimal" waveform, or class of waveforms, exists. For instance, it may be that CI algorithms should strive to approximate as closely as possible the temporal structure of neural responses that are produced by the normal ear in response to temporal fine structure. Conveying temporal fine-structure information via an envelope of a high-frequency carrier has been attempted in normal-hearing listeners using so-called "transposed stimuli" (TS). van der Par and Kohlrausch (1997) proposed modulating a high-frequency sinusoidal carrier with a half-wave rectified and lowpass-filtered low-frequency stimulus to pro-duce a temporal envelope that is intended to produce a response in the auditory nerve that resembles the response to the low-frequency stimulus itself. In this way the neural response to the low-frequency stimulus is "transposed" to a higher-frequency cochlear location. This approach has been used to show that some aspects of binaural hearing, such as binaural masking level differences, and interaural time resolution and lateralization for frequencies below 150 Hz, can be represented as well in high-frequency (4000 Hz and above) as in low-frequency cochlear regions (e.g., Henning, 1974; van der Par and Kohlrausch, 1997; Bernstein and Trahiotis, 2002). Most encouragingly, Long *et al.* (2006) used half-wave rectified (HWR) stimuli, mimicking the TS used in normal-hearing listeners, to show that CI users can benefit from information normally carried by the temporal fine structure to obtain a binaural masking release, thereby demonstrating the potential advantages of bilateral stimulation for spatial hearing and speech masking release.

It remains to be seen whether HWR stimuli can improve pitch perception for CI users. Results from acoustic studies in normal-hearing listeners showed that TS produced somewhat better pitch discrimination than sinusoidally amplitude modulated (SAM) stimuli (Oxenham *et al.*, 2004), although performance did not match that found using low-frequency pure tones. The aim of this study was to test pitch perception, quantified as modulation rate discrimination of high-rate pulse trains, by comparing the more traditional SAM with HWR stimuli. The hypothesis was that HWR stimuli may provide more accurate pitch perception by simulating more accurately the neural temporal patterns normally evoked by low-frequency pure tones.

II. METHODS

A. Subjects and implants

The subjects were seven postlingually deafened adults (two males, five females) with either a Clarion C-II or Hi-Res90K cochlear implant. Two subjects were bilaterally im-

---

[a] Author to whom correspondence should be addressed. Electronic mail: plumx@umn.edu
[b] Also affiliated with the Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455.
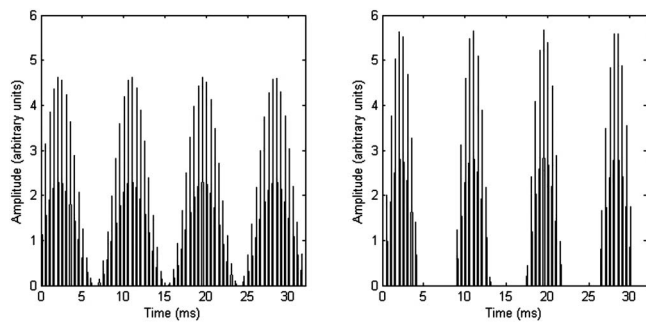
FIG. 1. Schematic diagram of the SAM stimulus (left panel) and HWR sinusoidal stimulus (right panel) with the same rms level. In this example, the pulse rate is 2000 Hz and the modulation rate is 115 Hz.

planted, and each ear was tested separately, resulting in a total of nine test ears. Full insertion of the electrode array (25 mm) was achieved in all cases. The age of the subjects ranged from 49 to 76 (mean age: 58 yr); the duration of CI use prior to the experiment ranged from 4 months to 5 years (mean duration: 2.8 yr); and the duration of deafness prior to implantation ranged from 1 to 28 years (mean duration: 14 yr). For the present study, stimulation was monopolar, with the active intracochlear electrode referenced to an electrode in the case of the internal receiver-stimulator. Electrodes were selected at three points across the array (apical, middle, and basal). In all cases, except for one subject's apical position (electrode 1), the test electrodes were 2, 8, and 14.

B. Stimuli and procedures

Experiments were controlled by a personal computer running custom programs written for the Bionic Ear Data Collection System (BEDCS; Advanced Bionics, 2003) and MATLAB (The Mathworks, Natick, MA). Stimuli were 300-ms trains of 32 μs/phase, cathodic-first biphasic pulses, presented at a rate of 2000 pulses per second. The amplitude of the pulses was modulated either by sinusoidal amplitude modulation (SAM condition) or by HWR sinusoidal modulation (HWR condition). Both the SAM and HWR conditions utilized 100% modulation depth with zero starting phase on each trial. Base modulation rates of 80, 115, 160, 230, and 320 Hz were tested. A schematic diagram of the two stimulus types is provided in Fig. 1. The levels of the stimuli were adjusted for each subject individually to correspond to either the 40% or 80% point of the subject's dynamic range (DR). The DR was defined as the range of currents (in μA) between absolute threshold (THS) and maximum acceptable loudness (MAL).

The THS and MAL were measured individually for each stimulus type at each modulation rate. The THS was measured using a two-interval, two-alternative forced-choice (2IFC/2AFC) procedure with a three-down, one-up adaptive tracking rule (Levitt, 1971). Correct-answer feedback was provided after each trial. The THS estimates from three to five tracks were averaged to obtain a final THS value for each subject. The MAL was measured using a one-up one-down adaptive tracking procedure in which the sound was presented, followed by the question "Was it too loud?." A subject's "no" and "yes" choices led to increases and de-

creases in signal level, respectively. The track terminated when the subject had responded that the intensity was too loud six times, and the average current level at the last six reversals was calculated. The MAL estimates from three tracks were averaged to obtain a final value of MAL for each subject.

Once the THS and MAL had been established for each subject, modulation rate discrimination thresholds were measured at each of the base modulation rates (80, 115, 160, 230, and 320 Hz) for both SAM and HWR modulations, using levels corresponding to 40% and 80% of the subject's DR. Discrimination thresholds were obtained using a 2IFC/2AFC procedure. No training was provided; however, 4 of the 7 subjects had extensive previous psychophysical testing experience. In order to reduce any potential loudness cues, the current level was roved (jittered) across intervals by $\pm 10\%$ DR for all subjects except one. For the remaining subject (also the subject for whom electrode 1 was tested), the jitter was set to $\pm 3\%$ DR, as this subject reported discomfort at jitter amounts greater than this. The two 300-ms stimulus intervals in each trial were separated by a silent interval of 500 ms. Within each of the two stimulus intervals, the pulse trains were modulated at rates of f and $f + \Delta f$, with the two modulation rates geometrically centered around the base modulation rate. The order of presentation of the two rates was selected randomly on each trial, and the subjects were instructed to select the interval containing the stimulus with the higher pitch. Correct-answer feedback was provided after each trial.

The value of Δf, expressed as a percentage of f, was varied adaptively using a three-down one-up procedure. Threshold was defined as the geometric mean value of Δf at the last six reversal points. Three such threshold estimates were averaged to obtain a single modulation rate discrimination threshold at each level for each electrode site. Conditions were tested in a blocked randomized order, with all conditions tested once before any were repeated. The order of presentation was selected randomly between subjects and between repetitions.

III. RESULTS AND DISCUSSION

A. THS and MAL

The THS and MAL measures for the apical, middle, and basal electrode locations, averaged across subjects, are plotted in terms of rms current as a function of modulation base rate in the left, middle, and right panels of Fig. 2, respectively. Generally, the THS values were similar for both HWR and SAM conditions, with average values (across modulation rate, electrode location, and subject) of 89.4 and 95.8 μA, respectively. The MAL values, in terms of rms current level, were higher for the SAM than for the HWR conditions. There was also a tendency for MAL values to increase from the apical to the basal electrode location, but the effect of modulation rate was less clear.

The fact that SAM and HWR produced similar THS values but different MAL values is in line with the results of Zhang and Zeng (1997). They used analog stimuli and found that detection thresholds were determined primarily by the

J. Acoust. Soc. Am., Vol. 127, No. 2, February 2010

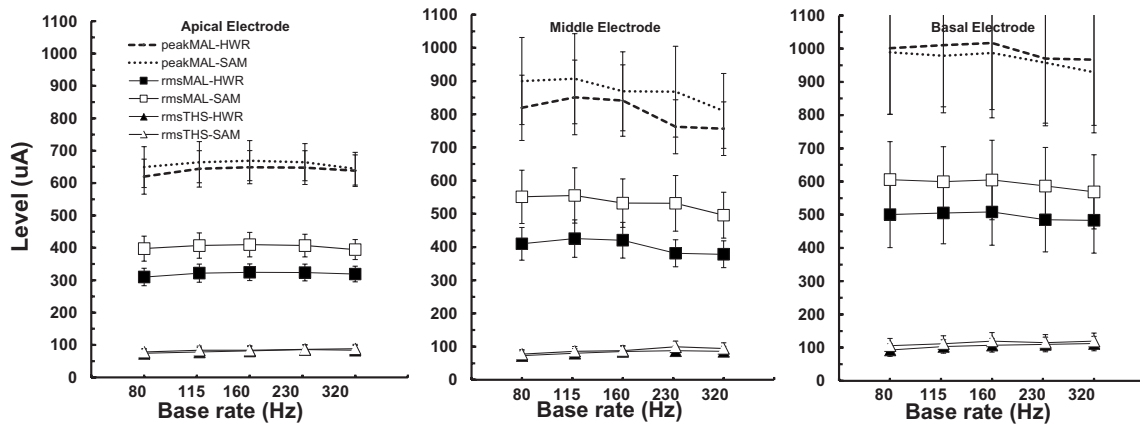Kreft *et al.*: Letters to the Editor   657

FIG. 2. THS and MAL as a function of base modulation rate averaged across subjects. Each panel represents the data for a different electrode location across the array. Open and filled symbols represent SAM and HWR stimuli, respectively. Triangles show THS values, squares show MAL values, and circles replot MAL values in terms of peak (rather than rms) amplitude. Error bars represent 1 s.e. of the mean.

rms amplitude of dynamic waveforms, whereas MAL was determined primarily by the peak amplitude. Using pulsatile stimuli, McKay and Henshall (2009) concluded that the relationship between modulated and unmodulated stimuli depends not only on level but also on pulse rate, in line with the model predictions of McKay et al. (2003). For a given rms amplitude, the peak amplitude of the HWR waveform is about 1.76 dB higher than that of the SAM waveform. This relationship is illustrated in Fig. 2 by the filled and open circles, which represent peak amplitudes at MAL for the HWR and SAM waveforms, respectively. Plotting MAL in terms of peak amplitude yields more similar results; in fact, the difference between HWR and SAM MAL values is not statistically significant when measured in terms of peak amplitude [$F(1,8)=1.912$, $p=0.204$].

B. Modulation rate discrimination

Figure 3 shows the just-noticeable change in modulation rate (Δf), expressed as a percentage of the lower rate (f), as a function of base modulation rate at current levels corresponding to 40% (squares) and 80% (triangles) of the subjects' dynamic range. Open and filled symbols represent thresholds using SAM and HWR waveforms, respectively.

As in Fig. 2, the left, middle, and right panels show data from the apical, middle, and basal electrodes, respectively. There was no obvious advantage for the HWR stimuli compared to the SAM stimuli. The size of the difference limens (DLs) for both HWR and SAM conditions was 6%–10% at the lowest base rate. This finding is consistent with other studies of modulation rate discrimination in cochlear-implant users (Geurts and Wouter, 2001; Baumann and Nobbe, 2004; Chatterjee and Peng, 2008; Landsberger, 2008).

In general, modulation rate DLs were lower (better) for the higher level stimulation (80% DR) than for the lower level (40% DR). There was also a clear trend for thresholds to increase (degrade) with increasing base modulation rate. In fact, thresholds were often so high for the highest two base rates (>30% of the base rate) that it seems unlikely that the stimuli evoked a reliable difference in pitch height. For instance, a 30% DL at a base rate of 230 Hz implies modulation rates for f and ($f+\Delta f$) of 198 and 267 Hz, respectively. It may be that such large differences were based on selecting the stimulus that produced the weaker, rather than higher, pitch percept. Modulation rate DLs were unaffected by electrode location. A repeated-measures analysis of variance formally tested all of the parameters associated with the
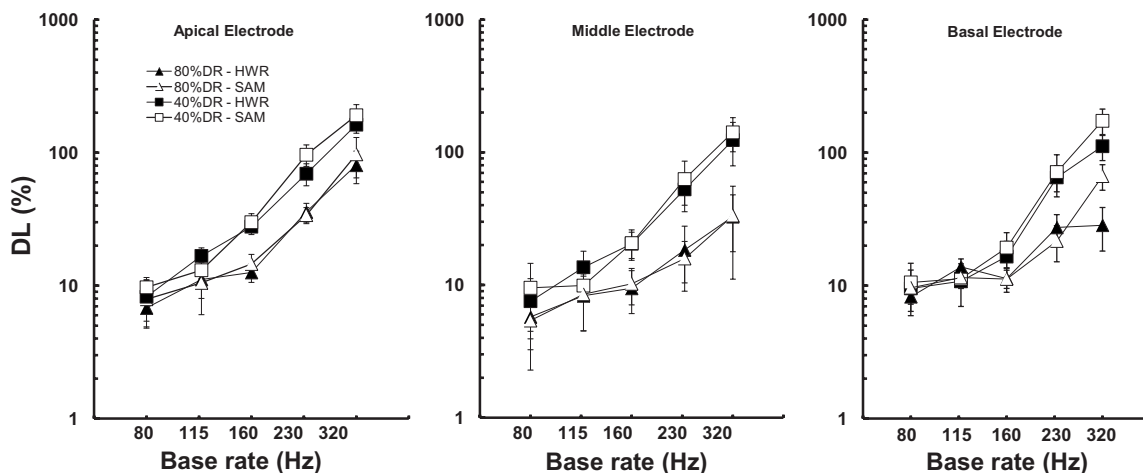

FIG. 3. Modulation rate DLs expressed as a percentage of the base rate averaged across subjects. Each panel represents the data for a different electrode location across the array. Error bars represent 1 s.e. of the mean.

modulation rate DLs. The effect of modulation rate was significant [$F(4,32)=40.864$, $p<0.001$], as was the effect of level [$F(1,8)=25.839$, $p=0.001$], whereas the effect of electrode location [$F(2,16)=2.732$, $p=0.127$] and waveform type [$F(1,8)=2.052$, $p=0.190$] was not. The only interaction term that was statistically significant was level $\times$ modulation rate [$F(4,32)=8.767$, $p<0.005$]. This pattern of results is in line with earlier studies of pulse-rate or modulation rate discrimination (for a review, see Moore and Carlyon, 2005).

The results do not support our initial hypothesis that HWR sinusoidal waveforms might provide more accurate modulation rate information than SAM waveforms. Our conclusion is consistent with that of Landsberger (2008), who tested a number of different waveform shapes and found that there were no significant differences in terms of rate discrimination abilities. The relatively poor performance in rate discrimination with temporal cues is in contrast to the relatively good performance in binaural masking level differences found by Long *et al.* (2006) also using the principle of TS. The fact that a benefit was found for spatial hearing but not for pitch is consistent with the idea that the temporal coding required for accurate pitch is different in nature from that required by the binaural system (e.g., Oxenham *et al.*, 2004). However, although the pattern of results is in the same direction as that found for normal-hearing listeners with acoustic stimulation, the details are somewhat different: normal-hearing listeners showed some benefit of TS over SAM in rate discrimination (Oxenham *et al.*, 2004), whereas CI users appear to show none. Taken together with previous results, the results suggest that accurate pitch discrimination in normal-hearing listeners requires place and/or place-time information (e.g., Loeb *et al.*, 1983; Shamma and Klein, 2000), and that accurate pitch coding in CI users may not be achieved by temporal patterns alone, but may require accurate spatial and temporal representation of sound along the tonotopic array.

ACKNOWLEDGMENTS

Advanced Bionics (**2003**). Bionic Ear Data Collection System, Version 1.16 User Manual.

Baumann, U., and Nobbe, A. (**2004**). "Pulse rate discrimination with deeply inserted electrode arrays," Hear. Res. **196**, 49–57.

Bernstein, L. R., and Trahiotis, C. (**2002**). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," J. Acoust. Soc. Am. **112**, 1026–1036.

Chatterjee, M., and Peng, S.-C. (**2008**). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition," Hear. Res. **235**, 143–156.

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

Galvin, J. J., III, Fu, Q. J., and Nogaki, G. (**2007**). "Melodic contour identification by cochlear implant listeners," Ear Hear. **28**, 302–319.

Geurts, L., and Wouters, J. (**2001**). "Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants," J. Acoust. Soc. Am. **109**, 713–726.

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., and Olszewski, C. (**2007**). "Accuracy of cochlear implant recipients on pitch perception, melody recognition and speech reception in noise," Ear Hear. **28**, 412–423.

Green, T., Faulkner, A., Rosen, S., and Macherey, O. (**2005**). "Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification," J. Acoust. Soc. Am. **118**, 375–385.

Henning, G. B. (**1974**). "Detectability of interaural delay in high-frequency complex waveforms," J. Acoust. Soc. Am. **55**, 84–90.

Landsberger, D. M. (**2008**). "Effects of modulation wave shape on modulation frequency discrimination with electric hearing," J. Acoust. Soc. Am. **124**, EL21–EL27.

Laneau, J., Wouters, J., and Moonen, M. (**2006**). "Improved music perception with explicit pitch coding in cochlear implants," Audiol. Neuro-Otol. **11**, 38–52.

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

Loeb, G. E., White, M. W., and Merzenich, M. M. (**1983**). "Spatial cross correlation: A proposed mechanism for acoustic pitch perception," Biol. Cybern. **47**, 149–163.

Long, C. J., Carlyon, R. P., Litovsky, R. Y., and Downs, D. H. (**2006**). "Binaural unmasking with bilateral cochlear implants," J. Assoc. Res. Otolaryngol. **7**, 352–360.

McDermott, H. J. (**2004**). "Music perception with cochlear implants: A review," Trends Amplif. **8**, 49–82.

McKay, C. M., and Henshall, K. R. (**2010**). "Amplitude modulation and loudness in cochlear implants," J. Assoc. Res. Otolaryngol. (in press).

McKay, C. M., Henshall, K. R., Farrell, R. J., and McDermott, H. J. (**2003**). "A practical method of predicting the loudness of complex electrical stimuli," J. Acoust. Soc. Am. **113**, 2054–2063.

Moore, B. C. J., and Carlyon, R. P. (**2005**). "Perception of pitch by people with cochlear hearing loss and by cochlear implant users," in *Pitch: Neural Coding and Perception*, edited by C. J. Plack, A. J. Oxenham, R. Fay, and A. N. Popper (Springer, New York).

Nelson, P. B., Jin, S.-H., Carney, A. E., and Nelson, D. A. (**2003**). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 961–968.

Oxenham, A. J., Bernstein, J. G. W., and Penagos, H. (**2004**). "Correct tonotopic representation is necessary for complex pitch perception," Proc. Natl. Acad. Sci. U.S.A. **101**, 1421–1425.

Qin, M. K., and Oxenham, A. J. (**2005**). "Effects of envelope-vocoder processing on F0 discrimination and concurrent-vowel identification," Ear Hear. **26**, 451–460.

Shamma, S., and Klein, D. (**2000**). "The case of the missing pitch templates: How harmonic templates emerge in the early auditory system," J. Acoust. Soc. Am. **107**, 2631–2644.

Stickney, G. S., Zeng, F. G., Litovsky, R., and Assmann, P. (**2004**). "Cochlear implant speech recognition with speech maskers," J. Acoust. Soc. Am. **116**, 1081–1091.

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.

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W., and McDermott, H. J. (**2005**). "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies," J. Acoust. Soc. Am. **117**, 3126–3138.

van der Par, Z., and Kohlraush, A. (**1997**). "A new approach to comparing binaural masking level differences at low and high frequencies," J. Acoust. Soc. Am. **101**, 1671–1680.

Zhang, C., and Zeng, F. G. (**1997**). "Loudness of dynamic stimuli in acoustic and electric hearing," J. Acoust. Soc. Am. **102**, 2925–2934.