Auditory enhancement and the role of spectral resolution in normal-hearing listeners and cochlear-implant users

Lei Feng^{a)} and Andrew J. Oxenham

Department of Psychology, University of Minnesota, N218 Elliott Hall, 75 East River Parkway, Minneapolis, Minnesota 55455, USA

(Received 23 March 2018; revised 25 June 2018; accepted 11 July 2018; published online 2 August 2018)

Detection of a target tone in a simultaneous multi-tone masker can be improved by preceding the stimulus with the masker alone. The mechanisms underlying this auditory enhancement effect may enable the efficient detection of new acoustic events and may help to produce perceptual constancy under varying acoustic conditions. Previous work in cochlear-implant (CI) users has suggested reduced or absent enhancement, due perhaps to poor spatial resolution in the cochlea. This study used a supra-threshold enhancement paradigm that in normal-hearing listeners results in large enhancement effects, exceeding 20 dB. Results from vocoder simulations using normal-hearing listeners showed that near-normal enhancement was observed if the simulated spread of excitation was limited to spectral slopes no shallower than 24 dB/oct. No significant enhancement was observed on average in CI users with their clinical monopolar stimulation strategy. The variability in enhancement between CI users, and between electrodes in a single CI user, could not be explained by the spread of excitation, as estimated from auditory nerve evoked potentials. Enhancement remained small, but did reach statistical significance, under the narrower partial-tripolar stimulation strategy. The results suggest that enhancement may be at least partially restored by improvements in the spatial resolution of current CIs. © 2018 Acoustical Society of America.

<https://doi.org/10.1121/1.5048414>

[JJL]

Pages: 552–566

I. INTRODUCTION

A target tone embedded in a simultaneous complex-tone masker will “pop out” from the mixture if the masker itself is presented as a precursor before the target-plus-masker mixture. This phenomenon is known as auditory enhancement. Enhancement has been quantified in normal-hearing (NH) listeners as the reduction in detection threshold of the target (Viemeister, 1980), or as the increase in forward masking produced by the target on a subsequent probe tone (Viemeister and Bacon, 1982; Byrne *et al.*, 2011). In many cases, the enhanced target can be salient enough to be perceived as a separate auditory entity with a distinct pitch (Hartmann and Goupell, 2006; Erviti *et al.*, 2011; Byrne *et al.*, 2013; Carcagno *et al.*, 2013). Similar effects can be observed when using noise bands as the precursor and/or masker (Viemeister *et al.*, 2013; Feng and Oxenham, 2015) and when the target is a more complex stimulus, such as the formants of a vowel (Summerfield *et al.*, 1987; Wang and Oxenham, 2016). These and other context effects may reflect the ability of the auditory system to normalize or “whiten” the representation of sound to improve coding efficiency (Barlow, 1961; Dean *et al.*, 2005), to sensitize the system to new events and important changes in the stimulus (Stilp *et al.*, 2010), and to generate perceptual constancy over a wide range of acoustic conditions, including different talkers, different room environments, and background noise (Ladefoged and Broadbent, 1957; Watkins, 1991).

Previous studies have proposed multiple mechanisms that may contribute to enhancement. One possible mechanism is the adaptation of neural responses to the maskers via the precursor, which makes the target response larger relative to the masker response (Viemeister, 1980). Another variation of this mechanism, known as adaptation of inhibition or suppression, involves the putative inhibition of the target by the masker components being adapted by the precursor, meaning that the target response is enhanced (Viemeister and Bacon, 1982; Wright *et al.*, 1993; Wang and Oxenham, 2016). Another type of explanation relies on perceptual grouping mechanisms (Carlyon, 1989), whereby the precursor and masker are perceptually grouped together, making the target stand out when introduced later, as an instantiation of the “old plus new” heuristic described by Bregman (1990). In addition, the precursor could provide other cues that help listeners to segregate the target from the subsequent masker, or simply provide a spectro-temporal cue when there is large uncertainty introduced by frequency roving from trial to trial (Richards *et al.*, 2004; Feng and Oxenham, 2015). The overall enhancement effect is likely to stem from multiple sources, originating in both the peripheral and central auditory system (Carcagno *et al.*, 2012; Feng and Oxenham, 2015).

Earlier studies have reported some enhancement in cochlear-implant (CI) users (Carlyon *et al.*, 2007; Goupell and Mostardi, 2012; Wang *et al.*, 2012). In cases where enhancement could be quantified using simultaneous masking or vowel identification, the size of the effect has been found to be either similar to that in NH listeners (Kreft and

^{a)}Electronic mail: fengl@umn.edu

Oxenham, 2017) or reduced in ways that could be explained by the poor spectral resolution of CIs (Carlyon *et al.*, 2007; Wang *et al.*, 2012). One exception appears to be enhancement under forward masking, where the enhancement was measured as the change in the amount of forward masking produced by the target on a following probe tone. Under these conditions, no enhancement was found in CI users (Kreft and Oxenham, 2017). The fact that some enhancement effects can be measured in CI users suggests that enhancement must be due at least in part to processes beyond the cochlea, as CI processing bypasses the cochlea and stimulates the auditory nerve directly.

The overlap of the electrical fields from neighboring electrodes is thought to be the primary limiter of spectral resolution in CI users, along with the limited number of electrodes and the potentially limited and uneven survival of spiral ganglion cells. If the reduced or absent enhancement reported in CI users is in part due to a loss of spectral resolution, the enhancement effect size observed should be related to the amount of channel interaction produced by current spread. Specifically, a larger enhancement effect should be related to better spectral resolution. In the current study, we measured auditory enhancement using targets that were presented at supra-threshold levels (Carcagno *et al.*, 2012; Feng and Oxenham, 2015). The listener's task was to discriminate the (spectral or place) pitch of the target under simultaneous masking. The rationale for choosing this behavioral task was threefold. First, the size of enhancement measured in such a configuration can be as large as 24 dB in NH listeners, meaning that it is likely to be of some importance in everyday listening situations. Second, the effect likely originates from multiple sources including frequency-specific adaptation, perceptual grouping, and spectrotemporal cueing (Feng and Oxenham, 2015), thereby providing an opportunity to investigate the effects of spectral resolution on different auditory processes. Third, the use of a discrimination task with a supra-threshold target arguably provides more ecological validity than a typical detection task. For instance, the identification and discrimination of spectral peaks is important for speech perception, as many speech sounds are distinguished by spectral cues, such as the frequencies of formant peaks (Stevens, 2000).

We assessed the link between auditory enhancement and spectral resolution in both CI users and NH listeners. In NH listeners, the link was tested by simulating different degrees of current spread using a noise-excited envelope vocoder. In CI users, the link was tested by comparing enhancement with measures of current spread using electrically evoked compound action potentials (ECAPs), and by comparing enhancement measured under two different stimulation modes: standard monopolar (MP) stimulation and the more focused partial tripolar (pTP) stimulation (Bierer, 2007). The questions posed by this study were as follows: (1) Is supra-threshold auditory enhancement reduced in CI users, relative to NH listeners? (2) To what extent is supra-threshold enhancement dependent on spectral resolution?

II. EXPERIMENT 1: EFFECTS OF SPECTRAL RESOLUTION ON AUDITORY ENHANCEMENT IN NORMAL-HEARING LISTENERS

A. Rationale

To test the effect of spectral resolution on supra-threshold enhancement, we presented NH listeners with stimuli that had been passed through a 16-channel noise-excited envelope vocoder, where the envelope extracted in each channel is used to modulate narrowband noises generated from bandpass synthesis filters. Different degrees of spectral resolution were simulated by varying the slopes of the reconstruction or synthesis filters from 6 to 24 dB/octave. The stimuli prior to vocoding were similar to those used by Feng and Oxenham (2015), and involved the presentation of a target tone, spectrally flanked by two masking tones, followed by a probe tone. The listeners' task was to judge whether or not the probe tone was presented at the same frequency as the target tone. Enhancement was defined as the threshold difference between the enhanced condition (ENH), when a precursor including the masker only was presented, and a control condition (CON), when the precursor included both the masker and target (Carcagno *et al.*, 2012; Feng and Oxenham, 2015). Because reduced spectral resolution will act to "fill in" the spectral gaps between target and masker components, we predicted that enhancement should decrease systematically with decreasing vocoder filter slopes. By using filter slopes of 6 and 12 dB/oct, which have been found to produce performance similar to that of CI users (Fu and Nogaki, 2005; Bingabr *et al.*, 2008; Oxenham and Kreft, 2014), the data provide a baseline with which to compare the performance of CI users in the same experimental paradigm.

B. Methods

1. Subjects

Twelve NH listeners (5 males, 7 females) participated in this experiment. All participants were between 19 and 30 years old and had audiometric pure-tone thresholds better than 20 dB hearing level (HL) at octave frequencies between 0.25 and 8 kHz. All participants provided written informed consent and were compensated for their time. All protocols were approved by the University of Minnesota Institutional Review Board.

2. Stimuli

The stimulus frequencies were selected to coincide with the center frequencies of the standard 16-channel clinical CI map from Advanced Bionics, as shown in Table I, with channel 1 corresponding to the most apical electrode and channel 16 to the most basal one. A schematic diagram of the stimuli is shown in Fig. 1. The target and masker components were of 100-ms duration, including 10-ms raised-cosine onset and offset ramps. The precursor, when present, was of 1000-ms duration, also including 10-ms ramps. The gap between the end of the precursor offset and the beginning of the target onset was 10 ms. The gap between the end of the masker and the beginning of the probe was set to

TABLE I. Standard 16-channel Advanced Bionics CI clinical map, showing the center frequency (CF) of each channel number.

Channel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
CF (Hz)	333	455	540	642	762	906	1076	1278	1518	1803	2142	2544	3022	3590	4264	6665

100 ms. The target center frequency was selected to coincide with the center frequency of filter corresponding to one of the channels 6 through 11 (corresponding to center frequencies of 906, 1076, 1278, 1518, 1803, or 2142 Hz; see Table I). The lower and upper maskers corresponded to the center frequencies of filters that were 4 channels below and above that of the target, respectively. The center frequency of the probe was either the same as that of the target (“same” trial) or higher or lower by one channel (“different” trial). As shown in Fig. 1, when a precursor was presented, it contained either just the masker components (ENH condition), or the masker and target components (CON condition). An independent white noise was generated and passed through a synthesis filter to stimulate each active filter channel. The synthesis filters were designed in MATLAB (Mathworks, Natick, MA) using the `fir2` function (Bingabr *et al.*, 2008) to produce slopes of 24, 12, and 6 dB/oct around the channel’s center frequency, with no flat passband region. The rationale for choosing those slopes is that Bingabr *et al.* (2008) used 24 dB/oct to simulate the current spread in MP mode in CI users, but other studies have found that the effective slopes in CI users may be as shallow as 6–8 dB/oct (Oxenham and Kreft, 2014). The level of the precursor and the maskers was set to 45 dB sound pressure level (SPL) per noise band. The level of the target and probe was adaptively varied, as described below.

3. Procedure

Participants were seated in a double-walled sound-attenuating booth. The stimuli were generated digitally using the AFC software package (Ewert, 2013) under MATLAB at a 48-kHz sampling rate, converted via an L22 soundcard

(LynxStudio, Costa Mesa, CA) with 24-bit resolution, and presented monaurally to the right ear through HD650 headphones (Sennheiser, Old Lyme, CT).

Before the main experiment, every participant was required to pass a basic pitch-discrimination training and screening session. In this session, the participants were presented with two consecutive bandpass noises, generated the same way as the target and probe in the main experiment (i.e., either at the same center frequency or with center frequencies differing by 1 channel), but without any precursor or masker. In each trial, participants were asked whether the two sounds had the same or different pitch. All participants had to obtain at least 80% correct (averaged across all trial types) for all three filter slopes (24, 12, and 6 dB/oct) to pass. All subjects passed the screening.

In the main experiment, each trial consisted of a single interval containing the combination of the masker and target, followed by the probe, in the simultaneous masker-only (MSK) condition. In the ENH and CON conditions, the masker was preceded by the precursor. The target and probe were always presented at the same level. The probe was either presented at the same center frequency as the target (50% probability) or at the next highest or next lowest center frequency (25% probability each). Listeners had to respond after each trial whether the target and probe were the same or different (Demany and Ramos, 2005; Feng and Oxenham, 2015). The target was randomly selected to be presented on one of the 6th–11th channels in every trial. Since the number of target and probe channels was limited, it was possible that listeners might use just the absolute frequency of the probe to perform the task, without reference to the target stimulus. In particular, when the probe was presented to channel 5 or 12, it was always different from the target. To avoid this

FIG. 1. (Color online) Schematic representation of the stimuli. Each interval includes a target (channel 8) embedded in two simultaneous maskers, located 4 channels lower (apically) and higher (basally) than the target, followed by a probe stimulus (grey). (a) The simultaneous masker condition without a precursor (MSK). (b) The enhanced condition (ENH), where a precursor including only the two maskers is presented before the target-plus-masker mixture. (c) The control condition (CON), where the target is included in the precursor. The probe could be presented in any one of the three possible positions shown in (a)–(c): same channel as the target or either of the two adjacent channels. The figures show the 250-ms precursor (as presented to the CI users) and 10-ms gap between the precursor and target.

potential confound, we added catch trials where the target and probe were set to be the same, either on channel 5 or 12, with equal probability as other possible combinations. Catch trials occurred with a total probability of 14%. They were not included in the adaptive procedure, and were not incorporated into threshold calculations.

The level of the target (and probe) was initially set to 65 dB SPL (i.e., 20 dB higher than the individual masker components, and so easily audible) and was varied adaptively using a 2-down 1-up adaptive tracking procedure that tracks the 70.7%-correct point on the psychometric function (Levitt, 1971). Because this is a “same-different” task, rather than a two-alternative forced-choice task (Macmillan and Creelman, 1991), the procedure assumes that the bias (i.e., the proclivity to respond “same” rather than “different”) did not vary substantially between subjects or between conditions. The initial step size of the tracking procedure was 5 dB; this was reduced to 2 dB after the first two reversals in the tracking procedure. The run was terminated after eight reversals and the threshold was computed as the average target level at the last six reversal points. No feedback was provided after each trial. All three conditions (ENH, CON, and MSK) were tested for each filter slope, for a total of nine test conditions. Each condition was tested four times for each participant, and the conditions were presented in a different random order for each repetition and each participant. Threshold was defined as the mean of the four runs for each test condition and participant. Before the experiment was run, all subjects were given a practice session, which included the ENH and MSK conditions, via a 32-channel noise vocoder and 96 dB/oct filter slopes.

C. Results

Thresholds in the three conditions (MSK, ENH, and CON) are shown in Fig. 2(a). Different symbol shades indicate the different filter slopes, which simulate different amounts of current spread. Lower thresholds were observed in the CON condition compared with the MSK condition. This is consistent with a previous study in NH listeners using inharmonic complex tones, which showed about 3 dB lower thresholds in the CON than in the MSK condition (Feng and Oxenham, 2015). This decrease does not seem to depend on

filter slope. In contrast, thresholds in the ENH condition depended on filter slope. With the 24 or 12 dB/oct current spread, thresholds in ENH condition were lower than those in either the MSK or CON condition, indicating enhancement. A within-subjects (repeated-measures) analysis of variance (ANOVA) was performed on the thresholds with factors of filter slope (6, 12, or 24 dB/oct) and precursor type (MSK, CON, or ENH). A Greenhouse-Geisser correction for lack of sphericity was applied where appropriate. The ANOVA showed a significant effect of filter slope [$F(1.43,15.8) = 8.45$, $p = 0.006$, partial $\eta^2 = 0.44$] and precursor type [$F(1.20,13.2) = 37.3$, $p < 0.001$, partial $\eta^2 = 0.77$]. There was also a significant interaction between filter slope and precursor type [$F(2.90,31.9) = 21.4$, $p < 0.001$, partial $\eta^2 = 0.66$]. Contrast analyses confirmed that thresholds in the MSK condition were significantly higher than in the ENH condition [$F(1,11) = 42.98$, $p < 0.001$, partial $\eta^2 = 0.80$] and CON condition [$F(1,11) = 33.67$, $p < 0.001$, partial $\eta^2 = 0.75$], and that thresholds in the CON condition were significantly higher than in the ENH condition [$F(1,11) = 28.91$, $p < 0.001$, partial $\eta^2 = 0.72$]. The significant interaction presumably reflects the fact that precursor type had a greater effect on thresholds when the filter slopes were steeper than when they were shallower.

The amount of enhancement, calculated as the difference in thresholds between the CON and ENH conditions (Carcagno *et al.*, 2012; Feng and Oxenham, 2015), is shown in Fig. 2(b) as a function of filter slope. When the filter slope was shallow (6 dB/oct slope), no significant enhancement was observed [paired t-test, $t(11) = 0.65$, $p = 0.53$]. However, enhancement was significant with the 12 and 24 dB/oct slopes [paired t-tests, 12 dB/oct: $t(11) = 3.39$, $p = 0.006$; 24 dB/oct: $t(11) = 6.68$, $p < 0.001$]. The amount of enhancement increased from 8 to 21 dB as the current spread reduced from 12 to 24 dB/oct. The amount of 21 dB observed with the 24 dB/oct slopes was comparable to the average of 24 dB enhancement found in NH listeners with inharmonic complex tones under similar conditions (Feng and Oxenham, 2015).

D. Discussion

The results from experiment 1 show that the amount of auditory enhancement decreases with decreasing spectral

FIG. 2. Averaged thresholds across 12 NH listeners in three different conditions and averaged auditory enhancement as a function of the filter slope. (a) The thresholds in the MSK, ENH, and CON conditions are plotted for each of the three filter slopes (6, 12, 24 dB/oct) in different shades. Error bars represent ± 1 standard error of the mean. (b) The average enhancement, defined as the difference between the thresholds measured in CON and ENH conditions, is plotted as a function of the filter slope. Error bars represent ± 1 standard error of the mean.

resolution. No significant enhancement was observed with the 6 dB/oct filter slopes. Interestingly, slopes of 24 dB/oct seem sufficient to restore these supra-threshold enhancement effects to near-normal levels, as determined by [Feng and Oxenham \(2015\)](#) for similar stimuli. This conclusion is in contrast to those based on the study of complex pitch perception, where much steeper slopes (>100 dB/oct) were necessary to obtain results comparable to normal hearing ([Mehta and Oxenham, 2017](#)). It seems that the spectral resolution necessary for normal enhancement effects is not as great as that needed to convey the F0 of complex tones via spectral peaks at harmonic frequencies.

As mentioned above, the spectral resolution of CIs has been found to be equivalent to vocoder filter slopes of somewhere between 8 and 24 dB/oct. Because this range spans the range from no enhancement to near-normal enhancement, it seems plausible that CI users might also exhibit supra-threshold enhancement effects, and that differences in spectral resolution, produced by individual differences or by different stimulation strategies, might lead to differences in enhancement. This prediction was tested in the remaining experiments, which employed CI users.

III. EXPERIMENT 2: AUDITORY ENHANCEMENT IN COCHLEAR-IMPLANT USERS AS A FUNCTION OF PRECURSOR DURATION AND PRECURSOR-MASKER GAP

A. Rationale

The primary aim of this experiment was to determine whether CI users experience auditory enhancement in conditions with a supra-threshold target. Previous studies have shown that auditory enhancement increases as the duration of the precursor increases and/or the gap duration between the precursor and masker decreases ([Viemeister, 1980](#); [Feng and Oxenham, 2015](#)). Measuring enhancement in multiple temporal configurations may help tease apart the contribution of different mechanisms. For instance, peripheral adaptation time constants are likely to be too short to account for effects that extend to durations of 1 s or more ([Harris and Dallos, 1979](#)). The secondary aim of this experiment was

therefore to determine the time course of any enhancement observed in CI users.

B. Methods

1. Subjects

Nine post-lingually deafened Advanced Bionics CI users participated in this study and were compensated for their time. Their information is provided in the first eight rows of Table II. All participants provided written informed consent. All protocols were approved by the Institutional Review Board of the University of Minnesota.

2. Stimuli

The stimuli were presented in the same sequence as those used for the NH listeners (Fig. 1). Instead of noise bands, the stimuli were presented as pulse trains via direct stimulation through a research interface. In the simultaneous MSK condition, each trial contained a target, randomly presented on one of the 6th through the 11th electrodes of the CI, embedded in two simultaneous maskers, which were located 4 electrodes higher and lower than the target [Fig. 1(a)]. This configuration ensured that only electrodes 2–15 were used in the experiment, since some CI users have the most distal electrode (1st or 16th) turned off. A probe followed the target-plus-masker mixture after a 200-ms silent gap. The silent gap was extended to make sure all CI users were able to hear the probe separately from the target-plus-masker mixture. The probe was either presented to the same electrode as the target or to one of the adjacent electrodes. As in experiment 1, the ENH condition included a precursor that was comprised of only the two maskers [Fig. 1(b)] and the CON condition included a precursor that was comprised of the target in addition to the two maskers [Fig. 1(c)].

All stimuli were delivered directly to the internal cochlear stimulator (ICS) system based on a clinical research platform, BEDCS (Advanced Bionics), and were presented as cathodic-leading biphasic pulse trains at a rate of 2000 pulses per second (pps) on each electrode. Each pulse was 32 μ s/phase. The three electrodes (two maskers and one target) were stimulated sequentially from apex to base for each

TABLE II. CI subject information. Presented in the columns from left to right are: Subject code, gender, age when tested for the present study, duration of implant use prior to the study, etiology of deafness, duration of hearing loss prior to implantation, the ICS type (HR90k = HiRes90K, HF90KA = HiRes90K Advantage), electrode type (HFII = HiFocus-II array, HF1j = HiFocus 1 J array, HFP = HiFocus array with positioner), and the number of experiments participated in the present study.

Subject code	Gender	Age (yr)	CI use (yr)	Etiology	Duration HL prior to implant (yr)	ICS Type	Electrode	Experiment Number
D02	F	64.9	13.1	Unknown	1	C-II	HF-II	Exp2,3,5
D10	F	61.5	12.9	Unknown	8	C-II	HF1J	Exp4,5
D19	F	55.4	10.7	Unknown	11	HF90k	HF1J	Exp2,3,4,5
D27	F	63.2	5.7	Otosclerosis	13	HF90k	HFP	Exp2
D28	F	65.8	11.8	Familial progressive SNHL	7	C-II	HF1J	Exp2,3,4,5
D33	M	75.7	2.3	Noise exposure; trauma	<1	HF90kA	HF1J	Exp2,3,4,5
D35	F	55.2	3.8	High fever	Unknown	HF90k	HF1J	Exp2
D38	F	32.8	1.6	Sudden	1	HF90kA	HF1J	Exp2
D39	M	67.5	5.9	Unknown	7	HF90k	HF1J	Exp2,3,4,5
D41	F	66.1	2.7	Familial progressive SNHL	41	HF90k	HF1J	Exp2,3,4,5
D42	M	60.2	2.3	Familial progressive SNHL	2	HF90k	HF1J	Exp4,5

pulse. All stimuli were presented in MP mode. The target, maskers, and probe were 100 ms in duration. The precursor duration was 62.5, 250, or 1000 ms, and the gap between the precursor and target was 10, 100, or 1000 ms. No onset or offset ramps were applied to the stimuli.

3. Procedure

The behavioral threshold (THS) and maximal acceptable level (MAL) of the 100-ms pulse train were measured for every electrode, EL 2–15, of each subject. The MAL of the 1000-ms pulse train was also measured to ensure that the longest precursor would not exceed the MAL when it was played at the same current level as the 100-ms masker. THS was measured using a three-interval, three-alternative forced-choice (3IFC/3AFC) procedure with a two-down, one-up adaptive tracking rule. Feedback was provided. The final THS was the average of two tracks. The MAL was measured using a one-up, one-down adaptive tracking procedure in which the participant was asked “Was the sound too loud?” A “no” or “yes” response led to an increase or decrease in signal level, respectively. The track terminated when the subject had responded six times that the sound was too loud, and the average current level at the last six reversals was calculated. The final MAL was also the average of two tracks.

Dynamic range (DR) is defined as the difference (in μA) between THS and MAL. In the main experiment, the current level for each electrode in the maskers and precursors was fixed at 20% DR of the 100-ms pulse train. The level of target was defined in units of dB, relative to its 20% DR level, similar to the target-to-masker ratio (TMR) in the study by Feng and Oxenham (2015). For instance, a target at 0 dB TMR is presented at 20% DR. The level of the target was initially set to 2.5 dB TMR and was varied adaptively following a two-down one-up rule that tracks the 70.7% correct point on the psychometric function (Levitt, 1971). No feedback was provided after each trial. The level of the probe was always the same as that of the target. Initially the target level was varied in steps of 0.5 dB. After two reversals in the direction of the adaptive tracking procedure, the step size was reduced to 0.2 dB. The run was terminated after eight reversals and the threshold was computed as the average target level at the last six reversal points of the tracking procedure. The estimated threshold represents the lowest target level for which the listeners could hear out the target and compare it to the probe at a performance level of 70.7% correct. As in experiment 1, catch trials with higher or lower target electrode numbers were added, but were not included in the adaptive procedure.

As in experiment 1, thresholds were measured in three conditions: no precursor (MSK), ENH, and the CON. In conditions with a precursor (ENH and CON), the nine combinations of three possible precursor durations and three durations of the gap between masker and precursor were tested. There were 19 conditions in total, including the MSK condition. Each condition was tested six times for each participant, and the conditions were presented in a different random order for each repetition and each participant.

Threshold was defined as the mean of the last four runs for each test condition and participant.

We ensured that all CI users were able to discriminate between adjacent electrodes by giving all subjects an initial training session, where only the target and probe were presented. Both sounds were presented at 40% DR with a $\pm 5\%$ DR roving. The subjects were asked whether the two sounds had the same or different pitches, and had to attain performance at a level corresponding to at least 80% correct to pass. All subjects passed the training. Feedback was provided in the training session.

C. Results

Thresholds in the different conditions are shown in Fig. 3(a), and the amount of enhancement, calculated as the dB difference in thresholds between the CON and ENH conditions, is shown in Fig. 3(b). Thresholds in the MSK condition seem higher than in the CON or ENH conditions, at least for the longest precursor durations and shortest gaps between the precursor and masker, suggesting some potential effects of cueing that are not related to enhancement. To test for the effect of the spectral notch in the precursor, and therefore enhancement, a three-way repeated-measures ANOVA was performed on the thresholds with factors of precursor type (ENH or CON), precursor duration (62.5, 250, or 1000 ms), and gap duration (10, 100, or 1000 ms), with a Greenhouse-Geisser correction for lack of sphericity, where appropriate. There was a significant main effect of precursor duration [$F(1.36, 10.9) = 18.7$, $p = 0.001$, partial $\eta^2 = 0.70$], as well as gap duration [$F(1.92, 15.3) = 14.5$, $p < 0.001$, partial $\eta^2 = 0.64$]. However, neither the main effect of precursor type nor any of its interactions reached significance [Main effect: $F(1, 8) = 1.25$, $p = 0.30$, partial $\eta^2 = 0.14$; interaction between precursor type and gap duration: $F(1.29, 10.34) = 0.38$, $p = 0.60$, partial $\eta^2 = 0.046$; interaction between precursor type and precursor duration: $F(1.24, 9.92) = 0.70$, $p = 0.45$, partial $\eta^2 = 0.08$; three-way interaction: $F(2.68, 21.4) = 0.31$, $p = 0.80$, partial $\eta^2 = 0.037$]. The interaction between gap duration and precursor duration was also not significant [$F(2.13, 17.0) = 2.29$, $p = 0.13$, partial $\eta^2 = 0.22$]. This lack of any significant difference in thresholds between the ENH and CON conditions suggests a general absence of supra-threshold enhancement effects in these CI users. Nevertheless, there did appear to be a trend towards some enhancement, particularly for the longest precursor and shortest gap [right panel of Fig. 3(b)]. There was also considerable individual variability, as shown by the individual data in Fig. 3(b).

D. Discussion

Although CI users did not show any enhancement on average from the spectral notch, there was a non-significant trend towards a benefit from the presence of a precursor, possibly through spectro-temporal cueing, as shown by the slightly lower thresholds in the CON than in the MSK conditions (for longer precursor durations and short gaps). This result is somewhat similar to the results from NH listeners with a precursor of a bandpass noise at long precursor

FIG. 3. (Color online) Time course of enhancement in CI users. (a) In each panel, the mean thresholds in the ENH and CON condition of all nine CI users are plotted as a function of gap duration for one precursor duration. The upper triangle represents the CON condition. The circle represents the ENH condition. The filled circle represents the threshold with no precursor (MSK condition). (b) Time course of enhancement of individual CI users (colored dotted lines) and the mean enhancement averaged across nine CI users (black solid line). Error bars represent ± 1 standard error.

(1000 ms) and short gap (10 ms) (Feng and Oxenham, 2015) and the results in experiment 1. However, for NH listeners, some reduction in thresholds with a bandpass-noise precursor was observed even with a short precursor (62.5 ms) or a long gap (1000 ms). It may be that the generally poorer spectral resolution experienced by CI users reduced the effectiveness of such cueing, particularly in the spectral domain.

Because the target was randomly selected to be one of the electrodes between 6 and 11 during our adaptive procedure, the measure of overall enhancement effect reflects the pooled responses across all the electrodes. Enhancement effects on some electrodes may have been negated by the responses from other electrodes, which might in part explain the lack of an overall enhancement effect measured in this experiment. It is known that spectral resolution can vary substantially between electrodes (Nelson *et al.*, 2011) due to various factors. For instance, the individual differences in electrode placement, intracochlear fibrous tissue, and bone growth (Kawano *et al.*, 1998; Ketten *et al.*, 1998; Finley *et al.*, 2008) may affect the current fields at the electrode-neural interface and hence the interactions between channels. Poor local neural survival near one electrode typically requires larger amounts of current to activate auditory nerve fibers located somewhere else, which leads to broader spread and reduced spectral resolution. The relationship between enhancement and spectral resolution at the level of

individual subjects and electrodes is pursued further in experiments 3 and 4.

IV. EXPERIMENT 3: EFFECTS OF ELECTRODE LOCATION ON ENHANCEMENT IN CI USERS

A. Rationale

In experiment 2, the thresholds were averaged across 6 possible target electrodes. The uneven activation pattern across the electrodes could create considerable variability in enhancement at different target electrodes (Shannon, 1983; Chatterjee and Shannon, 1998), potentially leading to an underestimate of enhancement on some electrodes. In this experiment, we examined enhancement on individual target electrodes.

B. Methods

The same basic method was used as in experiment 2, again with the target presented at random on one of electrodes 6 through 11. However, in this experiment, we organized the six target positions into three groups, electrodes 6 and 7 (apical), 8 and 9 (middle), and 10 and 11 (basal). The target level within each electrode group was tracked and varied independently in an interleaved tracking procedure, so that thresholds were estimated for each group separately within a

single run. We only tested the 1000-ms precursor with the 10-ms gap, as this was the condition most likely to produce enhancement (Fig. 3). The masker level was set at 40% DR at each electrode to test enhancement at a moderate sound level that is most likely used in everyday settings. Eight CI users, including six from experiment 2, were tested in the MSK, ENH, and CON conditions.

C. Results

The thresholds of individual CI users in the three conditions are shown in Fig. 4 as a function of electrode position. Most users (except D42 and D33) showed decreased thresholds in the ENH and CON conditions, relative to the MSK condition, on at least one electrode. To determine whether enhancement was present, a two-way repeated-measures ANOVA was performed on thresholds, with precursor type (ENH, CON) and electrode position (apical, middle, basal) as factors, including a Greenhouse-Geisser correction for lack of sphericity where appropriate. The ANOVA showed no significant effect of precursor type [$F(1,7)=3.61$, $p=0.099$, partial $\eta^2=0.34$]. Neither the effect of electrode position [$F(1.47,10.26)=0.006$, $p=0.98$, partial $\eta^2=0.001$], nor the interaction between precursor type and electrode position [$F(1.48,10.34)=0.075$, $p=0.88$, partial $\eta^2=0.011$], reached significance. Although there was no difference in thresholds at different electrode position on the population level, there was some variation within individual CI users (Fig. 4). Similarly, there was some hint of enhancement on some electrodes for some subjects (e.g., D02, D19, and D41). However, none of these differences reached significance at a population level, confirming the findings from

experiment 2 of no significant supra-threshold enhancement effects in CI users.

D. Discussion

Overall, no significant enhancement effects were observed, despite the examination of subjects and electrodes separately, and despite the fact that the stimuli were presented at a higher overall level. There were six CI users who were tested at both 20% DR (experiment 2) and 40% DR (experiment 3) masker levels in MSK and ENH condition. Using data from just these six listeners, a three-way repeated-measures ANOVA on thresholds with masker current level, precursor presence (MSK and ENH) and electrode position as factors showed a significant effect of masker current level [$F(1,5)=67.28$, $p<0.001$, partial $\eta^2=0.93$] but no significant two-way interaction between masker current level and other factors, or three-way interaction, $p>0.34$, suggesting that enhancement (if any) was not affected by the masker current level in the range tested.

Despite the overall lack of an enhancement effect, there were hints that some subjects experienced some enhancement on some electrodes. If these effects are due to better spectral resolution at those particular electrodes, then the measure of enhancement should correlate with measures of spectral (or spatial) resolution. Previous studies of spatial resolution in CIs have shown considerable between-subject and between-electrode variation, using measures such as forward-masked tuning curves, masking patterns, and spectral ripple detection or discrimination (Chatterjee and Shannon, 1998; Henry and Turner, 2003; Chatterjee *et al.*, 2006; Anderson *et al.*, 2011; Nelson *et al.*, 2011). The next experiment obtained physiological measures of spatial

FIG. 4. Thresholds of individual CI users plotted as a function of target electrode position in MSK, ENH, and CON conditions, respectively. Solid line with filled circles represents the MSK condition. Dashed line with open circles represents the ENH condition. Dashed line with open upper triangles represents the CON condition. Error bars represent the standard deviation of the mean averaged over four repeated measures.

resolution and compared them with individual measures of enhancement.

V. EXPERIMENT 4: SPATIAL MASKING PATTERNS AND ENHANCEMENT

A. Rationale

It is possible that the variation in enhancement within and between CI users in experiment 3 can be explained by an uneven pattern of spatial resolution within the cochlea. The spread of excitation within the cochlea in response to electrical pulse trains can be estimated with the ECAP (Cohen *et al.*, 2003; Abbas *et al.*, 2004). One method uses a probe fixed in electrode location while applying a masker stimulus to different electrodes across the array. The extent to which the presence of the masker affects the neural response to the probe provides an indication of the overlap in the neural population responding to both stimuli, which in turn yields an estimate of their spatial excitation patterns. In this experiment, we measured the spatial masking patterns of individual electrodes to investigate whether differing degrees of spread of excitation could explain some of the variance in the amount of enhancement measured behaviorally in experiment 3. A correlation between the behavioral measures of enhancement and the neural measures of channel interaction would support an explanation of the variance in enhancement based on differences in spatial resolution.

B. Methods

A forward-masking subtraction paradigm was used, in which the probe is fixed in level and location while the masker is systematically varied across the remaining electrodes in the array except for the recording electrode, which is typically two electrodes apical to the probe (Cohen *et al.*, 2003; Abbas *et al.*, 2004; Hughes and Abbas, 2006; Hughes and Stille, 2008). The masker and probe stimuli each consisted of a single cathodic-leading biphasic current pulse that was 32 μs /phase. The four-part ECAP stimulus set, including probe alone, masker plus probe, masker alone, and system signature, or baseline response without any stimulus, was repeated 60 to 100 times, depending on the signal-to-noise ratio of the ECAP waveforms. All stimuli were presented in MP mode with a masker-probe interval of 500 μs at a 20-Hz stimulation rate. The final waveform was averaged across all repetitions within a given subject and masker electrode location. We recorded from all eleven CI users who participated in behavioral measures but were only able to obtain reliable ECAP waveforms from seven of them (Table II). Variability in the accessibility of ECAPs has been reported in previous studies (Abbas *et al.*, 1999), and may be due to individual differences in the electrode-neural interface.

Stimulus levels were determined based on the behavioral dynamic range of the stimulus used to elicit the ECAP. To determine the dynamic range, ten sweeps of the ECAP stimulus were delivered in an ascending manner using a logarithmic step size (Hughes and Stille, 2009). Participants were asked to indicate when they first heard the sound (T) and when the sound was loud but not uncomfortable (C).

Behavioral dynamic range was defined as the difference (in μA) between C and T. Masker and probe were typically presented at 80% of the DR on each electrode. For D41, 90% DR was used instead because no reliable waveform was observed at 80% DR.

The ECAP data were read into a custom analysis program written in MATLAB. The waveform was derived from responses to the probe alone (S1), masker plus probe (S2), masker alone (S3), and system signature (S4): $\text{ECAP} = \text{S1} - \text{S2} + \text{S3} - \text{S4}$ (Abbas *et al.*, 2004). The ECAP amplitude was measured from the negative trough (N1) to the following positive peak/plateau (P2). ECAP amplitude was plotted as a function of masker electrode location to generate the spatial masking pattern. These patterns were measured for three target electrodes (EL 6, 8, 10) as well as their two maskers in experiment 3, for a total of seven electrodes (EL 2, 4, 6, 8, 10, 12, 14) for each CI subject.

C. Results

Figure 5 shows the examples of the spatial masking patterns of the electrodes corresponding to the target and the two maskers in the behavioral experiments above. Even though the maskers are both four electrodes away from the target, there is nevertheless a large degree of overlap between the spatial masking patterns of the target and the maskers, indicating possible channel interactions between them.

We used a channel interaction index (ChII) to quantify the area of overlap between maskers and the target electrode (Lu *et al.*, 2011). As shown in Fig. 6(a), the ChII was calculated as the ratio of the summed areas of masking electrodes that overlapped with the target electrode to the total cumulative areas of maskers and target, $\text{ChII} = (\text{XY} + \text{YZ})/(\text{X} + \text{Y} + \text{Z})$. Intuitively, a broader current spread should imply more masking of the target from the masker electrodes, which in turn should lead to higher thresholds under simultaneous masking. In addition, if enhancement is dependent on the degree of spectral resolution, then greater masker-target interaction should lead to less enhancement. To test these predictions, the ChII values (with each electrode measure from each subject treated as an independent measure) were correlated with the behavioral MSK thresholds [Fig. 6(b)] and with the amount of enhancement [Fig. 6(c)]. Neither correlation was significant [$r = 0.26$, $p = 0.26$ for thresholds in MSK and ChII; $r = -0.19$, $p = 0.41$ for enhancement and ChII]. This result suggests that the channel interaction pattern at the peripheral auditory system, as estimated by the ChII value, cannot explain the variability of enhancement within and across subjects, at least within the limited population size tested here.

D. Discussion

No correlation was observed between the variation in enhancement within and across subjects and the estimated spread of excitation at the peripheral auditory system. One possible explanation for the lack of correlation is that the channel interaction pattern at the peripheral auditory system alone cannot account for the variability of enhancement,

FIG. 5. Examples of SM patterns. (a)–(c) Each panel shows the SMs of a target at one of the three positions and its two maskers. The black arrow indicates the position of the target. Triangles represent the spatial masking pattern of the target electrode. The filled and open circles are the spatial masking patterns of the two maskers located 4 electrodes away from the target apically and basally.

since enhancement may be an accumulated effect generated from both peripheral and central parts of the auditory system. Another possible explanation is that even for electrode locations where the enhancement was observed, the size was fairly small, possibly due to the large spread of excitation with the MP stimulation. In the final experiment, to bypass individual differences of non-peripheral origins, and to directly manipulate spectral resolution (as was done in experiment 1 with the NH listeners), we compared enhancement under regular MP stimulation with that under pTP stimulation, which should lead to somewhat narrower excitation

patterns and improved spatial resolution (Litvak *et al.*, 2007; Bierer and Faulkner, 2010; Landsberger *et al.*, 2012).

VI. EXPERIMENT 5: AUDITORY ENHANCEMENT WITH PARTIAL TRIPOLAR STIMULATION

A. Rationale

With MP stimulation, an active electrode is stimulated and an extra-cochlear electrode is used as a ground, producing relatively broad current spread (Bierer and Middlebrooks, 2002; Bierer, 2007). With pTP stimulation, an active

FIG. 6. (a) Schematic diagram showing the calculation of channel interaction index (ChII) from the SMs of the target and maskers. (b) The ChII plotted as a function of the behavioral threshold in the no precursor (MSK) condition for all seven CI users on three target electrode positions. Different symbols represent different subjects. The coefficient of the regression line (solid line) was not significantly different from zero. (c) The ChII plotted as a function of the amount of enhancement found using MP stimulation. Again the regression coefficient was not significantly different from zero.

electrode is stimulated and a fraction (σ) of the return current goes to the two neighboring intracochlear electrodes with the rest ($1 - \sigma$) still returning to the extra-cochlear electrode. Previous studies have shown that current spread can be significantly reduced with pTP stimulation if σ is larger than 0.5 (Litvak *et al.*, 2007; Bierer and Faulkner, 2010; Landsberger *et al.*, 2012). In this experiment, we examined auditory enhancement with pTP stimulation, using the same paradigm as in experiment 3, and compared the enhancement measured with that produced by MP stimulation.

B. Methods

For pTP stimulation, the pulse trains were generated in the same way as for MP stimulation, except that two additional return electrodes were set to be the two neighboring electrodes, with a σ of 0.75. The THS and MAL for each electrode were measured with pTP stimulation using a fast method developed by Bierer *et al.* (2015) for both the 100-ms and 1000-ms pulse trains. The current level for each electrode in the precursor and masker was set to be 40% DR of the 100-ms pulse train. For some subjects, the current level reached the compliance limits before reaching MAL with pTP stimulation. In some such cases, the subjects were not able to reach 80% correct in the pitch discrimination training session with pTP stimulation at 40% DR. For these subjects, the current level with pTP stimulation was increased to the point where they could reach equal performance level in the electrode discrimination training comparing to that with MP stimulation at the current level of 40% DR. In other words, the current level was adjusted so that the CI users all had roughly equal baseline electrode-discrimination performance

when the target was presented in the absence of the maskers. The adjusted current levels for maskers with pTP stimulation varied between 50% DR and 80% DR across subjects. The target level started at 3 dB TMR and was varied adaptively. The initial step size of 0.5 dB was reduced to 0.2 dB after two reversals. We tested the MSK, ENH, and CON conditions with the 1000-ms precursor and 10-ms gap. Six of the nine CI users who participated in experiment 2 participated in this experiment. Two new CI users who participated in this experiment were also tested in all conditions with MP as well as pTP stimulation.

C. Results and discussion

Figure 7 shows the thresholds measured in the three conditions of MSK (no precursor), ENH (precursor without target), CON (precursor with target) at each target position for individual CI users with pTP stimulation. In general, the thresholds in the ENH condition were often, but not always, lower than in the MSK and CON conditions, indicating some enhancement. Some subjects even attained thresholds at negative TMRs in the ENH condition (D02, D19, D10). Some subjects also appeared to benefit from a precursor including the target (CON) on some or all target locations (D39, D10, D02, and D33). To establish whether enhancement was observed, a two-way repeated-measures ANOVA was performed on the masked thresholds with precursor type (ENH, CON) and electrode position (EL6,7, EL8,9, EL10,11) as factors. The ANOVA showed a significant effect of precursor type [$F(1,7) = 7.23$, $p = 0.031$, partial $\eta^2 = 0.51$], which suggests a small but significant enhancement effect with pTP stimulation. The effect of electrode

FIG. 7. Thresholds using pTP stimulation for individual CI users on three target electrode positions in MSK, ENH, and CON conditions. The solid line with filled circles represents the MSK condition. The dashed line with open circles represents the ENH condition. The dashed line with open triangles represents the CON condition. Error bars are the standard deviation (SD) of the mean thresholds across four repeated measures.

position was not significant [$F(1.97,13.79) = 2.45$, $p = 0.12$, partial $\eta^2 = 0.26$], and there was no interaction between precursor type and electrode position [$F(1.47,10.31) = 2.10$, $p = 0.18$, partial $\eta^2 = 0.23$]. The enhancement of individual users, i.e., the threshold difference between the CON and ENH condition, is shown as circles in Fig. 8.

Comparing the results from this experiment with those from experiment 3, there was a tendency towards greater enhancement with the pTP stimulation (0.60 dB on average across all electrode position and all CI users) than with MP stimulation (0.43 dB on average). However, the difference in enhancement between pTP and MP did not reach significance. A three-way repeated-measures ANOVA was performed on the masked thresholds with stimulation mode (MP, pTP), precursor type (ENH, CON) and electrode position as factors. The effect of precursor type was significant [$F(1,7) = 11.37$, $p = 0.012$, partial $\eta^2 = 0.62$], but neither the effect of stimulation mode [$F(1,7) = 3.79$, $p = 0.092$, partial $\eta^2 = 0.35$] nor the interaction between stimulation mode and precursor type [$F(1,7) = 0.34$, $p = 0.58$, partial $\eta^2 = 0.046$] was significant. The effect of electrode position was not significant [$F(1.46,10.19) = 1.51$, $p = 0.26$, partial $\eta^2 = 0.18$], and none of the other interactions was significant ($p > 0.53$).

Similar to MP stimulation, the variability of enhancement within subjects can be as large as that across subjects with pTP stimulation. For instance, D02 showed enhancement only on apical and middle targets, whereas D33 showed no enhancement on middle target but about 1 dB enhancement on apical and basal target. One subject (D28) showed no enhancement at all. Nevertheless, on average, with pTP stimulation the effect of enhancement did reach significance for the eight CI subjects tested here, in contrast to what was found in experiments 2 and 3 under MP stimulation. However, the effect was small, and the difference in enhancement between MP and pTP did not reach significance when tested directly.

VII. GENERAL DISCUSSION

A. Summary of results

The effects of spectral resolution on supra-threshold auditory enhancement were investigated in CI users and in noise-vocoder simulations of CI processing in NH listeners. Results from NH listeners showed that enhancement effects were at near-normal levels of 20 dB when spectral resolution was limited to 16 channels with filter slopes of 24 dB/oct, but decreased rapidly to near zero when spectral resolution was further degraded by shallower filter slopes of 6 dB/oct (experiment 1). No significant enhancement was found with CI users under MP stimulation, although large individual differences were observed between subjects and between electrodes (experiments 2 and 3). No significant relationships were found in the CI users between spatial resolution, as estimated via ECAP, and measures of enhancement (experiment 4). Finally, the use of more focused electrical fields, via pTP stimulation, resulted in a small but significant enhancement effect in the CI users, although the difference between MP and pTP did not reach significance (experiment 5).

B. Comparison with earlier studies

A study of auditory enhancement under simultaneous masking by Kreft and Oxenham (2017) found significant enhancement in CI users using MP stimulation. In their paradigm, as in the original study of Viemeister (1980), detection thresholds were measured for a target in the presence of a spectrally notched masker. One difference in the paradigms is that the previous studies measured enhancement of the target at threshold levels, whereas the current study measured enhancement at supra-threshold levels, and required a discrimination judgment, rather than just detection of the target. Some enhancement was also reported in CI users listening through their speech processors with a vowel discrimination

FIG. 8. Enhancement of individual CI users at three target positions with MP and pTP stimulation, respectively. Enhancement is defined as the difference in thresholds in CON and ENH conditions. The dashed line with filled circle represents enhancement with pTP stimulation. The solid line with filled upper triangle represents enhancement with MP stimulation.

task, which required identification of spectral peaks (Wang *et al.*, 2012), although the amount of enhancement was reduced relative to that found in NH listeners, presumably due to poorer spectral resolution.

Another difference between the task used in the present study and those from earlier studies is that performance in the present task depends on “hearing out” the target from the masker, whereas all the other tasks can in principle be performed by perceiving (and detecting changes in) the entire target-plus-masker complex as a single object. It may be that greater spectral resolution (equivalent to at least 24 dB/oct) is required for perceptual object segregation. Although significant enhancement was measured with pTP stimulation, the average effect size was about 0.6 dB, or 6% of the overall dynamic range of CI users. If the dynamic range of normal-hearing listeners is considered to be about 100 dB (Wang *et al.*, 2015, 2016), the average enhancement of 6% DR observed in CI users would correspond to 6 dB which is still much smaller than the 24 dB found in NH listeners with inharmonic complex tones (Feng and Oxenham, 2015).

C. Implications for underlying mechanisms

Because no enhancement was observed in the main experiment involving CI users with MP stimulation (experiment 2), it was not possible to conclude anything regarding the time constants involved in enhancement for CI users. Therefore, the results from that experiment do not provide any indication regarding the peripheral or central locus of the effect. The lack of clear correlations between measures of spectral resolution and enhancement in individual CI users also precludes any strong conclusion about the role of spectral resolution in CI users. However, the fact that spectral resolution played a critical role in the data from NH listeners suggests that it may also underlie the lack of enhancement effects in CI users. Indeed, if spectral resolution in CI users is approximated by vocoder processing with slopes of around 8 dB/oct, as suggested by Oxenham and Kreft (2014), then it is perhaps not surprising that CI users under MP stimulation did not show any enhancement, given the fact that NH listeners also showed no enhancement with similar (6 dB/oct) slopes. The fact that enhancement began to emerge under pTP stimulation is a further indication that CI users’ lack of enhancement may be explained by poor spectral resolution.

D. Implications for speech perception and processing

The enhancement effects observed in the paradigm used here can be very large—exceeding 20 dB in NH listeners (Feng and Oxenham, 2015). The loss of such a large perceptual effect may have important consequences for speech perception in everyday environments. For instance, if enhancement increases the salience of new objects in the acoustic environment, then its loss implies that it may be more difficult to detect and recognize new sounds and new voices in an ongoing acoustic background (Gregg and Snyder, 2012). Similarly, to the extent that enhancement allows the auditory system to adjust to new acoustic environments, and hence facilitate

perceptual constancy, this ability will be degraded in CI users. Although a significant effect was observed in CI users with pTP stimulation, the effect size might be too small to provide meaningful sensitivity adjustment.

Marrufo-Pérez *et al.* (2018) found that CI users showed adaptation to noise, which facilitated speech recognition if the noise started earlier than the speech sound. The amount of adaptation to noise was similar to that found in NH listeners. This type of adaptation can be considered a “within-channel” effect, as the noise and speech cover the same spectral region. This form of adaptation may reflect the adjustment of neural dynamic range to the sound level distribution (Dean *et al.*, 2005; Rabinowitz *et al.*, 2011), but does not require frequency selectivity. In our study, the target was presented in a different spectral region from the masker, so that the enhancement effect is defined as the difference between the effects found with and without the target component in the precursor. Such frequency-selective effects may also rely on adaptation but result in an enhancement of local spectral contrast. This type of adaptation could help in the detection of changes in ongoing stimuli, and may affect the perception of vowels or consonants, as shown in earlier studies of spectral contrast effects in speech (Watkins, 1991; Holt, 2006).

One encouraging aspect here is that the kind of enhancement measured in the present study could, in principle, be reintroduced into CIs via channel-specific automatic gain control. Gain in any specific channel could be adjusted based on the preceding level of stimulus within that channel, so that if a spectral gap was placed within that channel, then its gain would be increased, relative to that of the surrounding channels. It remains to be determined whether such a scheme would restore enhancement, as measured here, and whether it would result in improved speech perception in complex settings.

ACKNOWLEDGMENTS

This research was supported by National Institute of Health (NIDCD Grant No. R01 DC012262). The authors thank Heather Kreft for assistance with subject recruitment and Dr. Michelle Hughes for advice on ECAP recordings. The authors also wish to extend thanks to the subjects who participated in this study.

- Abbas, P. J., Brown, C. J., Shallop, J. K., Firszt, J. B., Hughes, M. L., Hong, S. H., and Staller, S. J. (1999). “Summary of results using the nucleus CI24M implant to record the electrically evoked compound action potential,” *Ear Hear.* **20**, 45–59.
- Abbas, P. J., Hughes, M. L., Brown, C. J., Miller, C. A., and South, H. (2004). “Channel interaction in cochlear implant users evaluated using the electrically evoked compound action potential,” *Audiol. Neurotol.* **9**, 203–213.
- Anderson, E. S., Nelson, D. A., Kreft, H., Nelson, P. B., and Oxenham, A. J. (2011). “Comparing spatial tuning curves, spectral ripple resolution, and speech perception in cochlear implant users,” *J. Acoust. Soc. Am.* **130**, 364–375.
- Barlow, H. B. (1961). “Possible principles underlying the transformation of sensory messages,” in *Sensory Communication*, edited by W. Rosenblith (MIT Press, Cambridge, MA), pp. 217–234.
- Bierer, J. A. (2007). “Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration,” *J. Acoust. Soc. Am.* **121**, 1642–1653.

- Bierer, J. A., Bierer, S. M., Kreft, H. A., and Oxenham, A. J. (2015). "A fast method for measuring psychophysical thresholds across the cochlear implant array," *Trends Hear* **19**, 1–12.
- Bierer, J. A., and Faulkner, K. F. (2010). "Identifying cochlear implant channels with poor electrode-neuron interface: Partial tripolar, single-channel thresholds and psychophysical tuning curves," *Ear Hear.* **31**, 247–258.
- Bierer, J. A., and Middlebrooks, J. C. (2002). "Auditory cortical images of cochlear-implant stimuli: Dependence on electrode configuration," *J. Neurophysiol.* **87**, 478–492.
- Bingabr, M., Espinoza-Varas, B., and Loizou, P. C. (2008). "Simulating the effect of spread of excitation in cochlear implants," *Hear. Res.* **241**, 73–79.
- Bregman, A. S. (1990). *Auditory Scene Analysis* (MIT Press, Cambridge, MA).
- Byrne, A. J., Stellmack, M. A., and Viemeister, N. F. (2011). "The enhancement effect: Evidence for adaptation of inhibition using a binaural centering task," *J. Acoust. Soc. Am.* **129**, 2088–2094.
- Byrne, A. J., Stellmack, M. A., and Viemeister, N. F. (2013). "The salience of enhanced components within inharmonic complexes," *J. Acoust. Soc. Am.* **134**, 2631–2634.
- Carcagno, S., Semal, C., and Demany, L. (2012). "Auditory enhancement of increments in spectral amplitude stems from more than one source," *J. Assoc. Res. Otolaryngol.* **13**, 693–702.
- Carcagno, S., Semal, C., and Demany, L. (2013). "No need for templates in the auditory enhancement effect," *PLoS One* **8**, e67874.
- Carlyon, R. P. (1989). "Changes in the masked thresholds of brief tones produced by prior bursts of noise," *Hear. Res.* **41**, 223–235.
- Carlyon, R. P., Long, C. J., Deeks, J. M., and McKay, C. M. (2007). "Concurrent sound segregation in electric and acoustic hearing," *J. Assoc. Res. Otolaryngol.* **8**, 119–133.
- Chatterjee, M., Galvin, J. J., III, Fu, Q. J., and Shannon, R. V. (2006). "Effects of stimulation mode, level and location on forward-masked excitation patterns in cochlear implant patients," *J. Assoc. Res. Otolaryngol.* **7**, 15–25.
- Chatterjee, M., and Shannon, R. V. (1998). "Forward masked excitation patterns in multielectrode electrical stimulation," *J. Acoust. Soc. Am.* **103**, 2565–2572.
- Cohen, L. T., Richardson, L. M., Saunders, E., and Cowan, R. S. (2003). "Spatial spread of neural excitation in cochlear implant recipients: Comparison of improved ECAP method and psychophysical forward masking," *Hear. Res.* **179**, 72–87.
- Dean, I., Harper, N. S., and McAlpine, D. (2005). "Neural population coding of sound level adapts to stimulus statistics," *Nat. Neurosci.* **8**, 1684–1689.
- Demany, L., and Ramos, C. (2005). "On the binding of successive sounds: Perceiving shifts in nonperceived pitches," *J. Acoust. Soc. Am.* **117**, 833–841.
- Erviti, M., Semal, C., and Demany, L. (2011). "Enhancing a tone by shifting its frequency or intensity," *J. Acoust. Soc. Am.* **129**, 3837–3845.
- Ewert, S. D. (2013). "A modular framework for running psychoacoustic experiments and computational perception models," in *International Conference on Acoustics AIA-DAGA*, Merano, Italy, pp. 1326–1329.
- Feng, L., and Oxenham, A. J. (2015). "New perspectives on the measurement and time course of auditory enhancement," *J. Exp. Psychol. Hum. Percept. Perform.* **41**, 1696–1708.
- Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., Hullar, T. E., and Skinner, M. W. (2008). "Role of electrode placement as a contributor to variability in cochlear implant outcomes," *Otol. Neurotol.* **29**, 920–928.
- Fu, Q. J., and Nogaki, G. (2005). "Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing," *J. Assoc. Res. Otolaryngol.* **6**, 19–27.
- Goupell, M. J., and Mostardi, M. J. (2012). "Evidence of the enhancement effect in electrical stimulation via electrode matching (L)," *J. Acoust. Soc. Am.* **131**, 1007–1010.
- Gregg, M. K., and Snyder, J. S. (2012). "Enhanced sensory processing accompanies successful detection of change for real-world sounds," *NeuroImage* **62**, 113–119.
- Harris, D. M., and Dallos, P. (1979). "Forward masking of auditory nerve fiber responses," *J. Neurophysiol.* **42**, 1083–1107.
- Hartmann, W. M., and Goupell, M. J. (2006). "Enhancing and unmasking the harmonics of a complex tone," *J. Acoust. Soc. Am.* **120**, 2142–2157.
- 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.
- Holt, L. L. (2006). "The mean matters: Effects of statistically defined non-speech spectral distributions on speech categorization," *J. Acoust. Soc. Am.* **120**, 2801–2817.
- Hughes, M. L., and Abbas, P. J. (2006). "Electrophysiologic channel interaction, electrode pitch ranking, and behavioral threshold in straight versus perimodular cochlear implant electrode arrays," *J. Acoust. Soc. Am.* **119**, 1538–1547.
- Hughes, M. L., and Stille, L. J. (2008). "Psychophysical versus physiological spatial forward masking and the relation to speech perception in cochlear implants," *Ear Hear.* **29**, 435–452.
- Hughes, M. L., and Stille, L. J. (2009). "Psychophysical and physiological measures of electrical-field interaction in cochlear implants," *J. Acoust. Soc. Am.* **125**, 247–260.
- Kawano, A., Seldon, H. L., Clark, G. M., Ramsden, R. T., and Raine, C. H. (1998). "Intracochlear factors contributing to psychophysical percepts following cochlear implantation," *Acta Otolaryngol.* **118**, 313–326.
- Ketten, D. R., Skinner, M. W., Wang, G., Vannier, M. W., Gates, G. A., and Neely, J. G. (1998). "In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays," *Ann. Otol. Rhinol. Laryngol. Suppl.* **175**, 1–16.
- Kreft, H. A., and Oxenham, A. J. (2017). "Auditory enhancement in cochlear-implant users under simultaneous and forward masking," *J. Assoc. Res. Otolaryngol.* **18**, 483–493.
- Ladefoged, P., and Broadbent, D. E. (1957). "Information conveyed by vowels," *J. Acoust. Soc. Am.* **29**, 98–104.
- Landsberger, D. M., Padilla, M., and Srinivasan, A. G. (2012). "Reducing current spread using current focusing in cochlear implant users," *Hear. Res.* **284**, 16–24.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Litvak, L. M., Spahr, A. J., and Emadi, G. (2007). "Loudness growth observed under partially tripolar stimulation: Model and data from cochlear implant listeners," *J. Acoust. Soc. Am.* **122**, 967–981.
- Lu, T., Litovsky, R., and Zeng, F. G. (2011). "Binaural unmasking with multiple adjacent masking electrodes in bilateral cochlear implant users," *J. Acoust. Soc. Am.* **129**, 3934–3945.
- Macmillan, N. A., and Creelman, C. D. (1991). *Detection Theory: A User's Guide* (Cambridge University Press, Cambridge, England).
- Marruf-Perez, M. I., Eustaquio-Martin, A., and Lopez-Poveda, E. A. (2018). "Adaptation to noise in human speech recognition unrelated to the medial olivocochlear reflex," *J. Neurosci.* **38**, 4138–4145.
- Mehta, A. H., and Oxenham, A. J. (2017). "Vocoder simulations explain complex pitch perception limitations experienced by cochlear implant users," *J. Assoc. Res. Otolaryngol.* **18**, 789–802.
- Nelson, D. A., Kreft, H. A., Anderson, E. S., and Donaldson, G. S. (2011). "Spatial tuning curves from apical, middle, and basal electrodes in cochlear implant users," *J. Acoust. Soc. Am.* **129**, 3916–3933.
- Oxenham, A. J., and Kreft, H. A. (2014). "Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing," *Trends Hear* **18**, 1–14.
- Rabinowitz, N. C., Willmore, B. D., Schnupp, J. W., and King, A. J. (2011). "Contrast gain control in auditory cortex," *Neuron* **70**, 1178–1191.
- Richards, V. M., Huang, R., and Kidd, G., Jr. (2004). "Masker-first advantage for cues in informational masking," *J. Acoust. Soc. Am.* **116**, 2278–2288.
- Shannon, R. V. (1983). "Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction," *Hear. Res.* **12**, 1–16.
- Stevens, K. N. (2000). *Acoustic Phonetics* (MIT Press, Cambridge, MA).
- Stilp, C. E., Alexander, J. M., Kieffe, M., and Kluender, K. R. (2010). "Auditory color constancy: Calibration to reliable spectral properties across nonspeech context and targets," *Atten. Percept. Psychophys.* **72**, 470–480.
- Summerfield, Q., Sidwell, A., and Nelson, T. (1987). "Auditory enhancement of changes in spectral amplitude," *J. Acoust. Soc. Am.* **81**, 700–708.
- Viemeister, N. F. (1980). "Adaptation of masking," in *Psychophysical, Physiological and Behavioural Studies in Hearing*, edited by G. van den Brink and F. A. Bilsen (Delft University Press, Delft, the Netherlands), pp. 190–198.
- Viemeister, N. F., and Bacon, S. P. (1982). "Forward masking by enhanced components in harmonic complexes," *J. Acoust. Soc. Am.* **71**, 1502–1507.
- Viemeister, N. F., Byrne, A. J., and Stellmack, M. A. (2013). "Spectral and level effects in auditory signal enhancement," *Adv. Exp. Med. Biol.* **787**, 167–174.

- Wang, N., Kreft, H., and Oxenham, A. J. (2012). "Vowel enhancement effects in cochlear-implant users," *J. Acoust. Soc. Am.* **131**, EL421–EL426.
- Wang, N., Kreft, H., and Oxenham, A. J. (2016). "Induced loudness reduction and enhancement in acoustic and electric hearing," *J. Assoc. Res. Otolaryngol.* **17**, 383–391.
- Wang, N., Kreft, H. A., and Oxenham, A. J. (2015). "Loudness context effects in normal-hearing listeners and cochlear-implant users," *J. Assoc. Res. Otolaryngol.* **16**, 535–545.
- Wang, N., and Oxenham, A. J. (2016). "Effects of auditory enhancement on the loudness of masker and target components," *Hear. Res.* **333**, 150–156.
- Watkins, A. J. (1991). "Central auditory mechanisms of perceptual compensation for spectral-envelope distortion," *J. Acoust. Soc. Am.* **90**, 2942–2955.
- Wright, B. A., McFadden, D., and Champlin, C. A. (1993). "Adaptation of suppression as an explanation of enhancement effects," *J. Acoust. Soc. Am.* **94**, 72–82.