Sensitivity to interaural envelope correlation changes in bilateral cochlear-implant users

Matthew J. Goupell[a]
Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742

Ruth Y. Litovsky
Waisman Center, University of Wisconsin, 1500 Highland Avenue, Madison, Wisconsin 53705

Provision of bilateral cochlear implants (CIs) to people who are deaf is partially justified by improved abilities to understand speech in noise when comparing bilateral vs unilateral listening conditions. However, bilateral CI listeners generally show only monaural head shadow with little improvement in speech understanding due to binaural unmasking. Sensitivity to change in interaural envelope correlation, which is related to binaural speech unmasking, was investigated. Bilateral CI users were tested with bilaterally synchronized processors at single, pitch-matched electrode pairs. First, binaural masking level differences (BMLDs) were measured using 1000 pulse-per-second (pps) carriers, yielding BMLDs of 11.1 ± 6.5 and 8.5 ± 4.2 dB for 10- and 50-Hz bandwidth masking noises, respectively. Second, envelope correlation change just-noticeable differences (JNDs) were measured. Stimuli presented at 1000 pps yielded lower JNDs than those presented at 100 pps. Furthermore, perfectly correlated reference stimuli produced lower JNDs than uncorrelated references, and uncorrelated references generally produced immeasurable JNDs. About 25% of JNDs measured in the CI listeners were in the range of JNDs observed in normal-hearing listeners presented CI simulations. In conclusion, CI listeners can perceive changes in interaural envelope correlation, but the poor performance may be a major limiting factor in binaural unmasking tested to date in realistic listening environments. © 2015 Acoustical Society of America.
[http://dx.doi.org/10.1121/1.4904491]

I. INTRODUCTION

Binaural hearing provides a vastly improved ability to localize sound sources and to understand speech in noisy environments compared to monaural hearing. This is due to the fact that the binaural system is exquisitely sensitive to differences in the physical properties of signals that reach a listener's ears. These differences between the signals at the two ears manifest themselves as interaural time differences (ITDs, the long-term average difference in the time that it physically takes a sound to travel to each ear), interaural level differences (ILDs, the long-term average difference in level produced by the acoustic shadow that the head introduces), and interaural decorrelation. The latter arises when the signals between the two ears are different in ways that are not captured when the long-term ITDs or ILDs are computed. Interaural decorrelation can be mathematically characterized by the normalized interaural cross-correlation function, which is defined as

$$\rho(\tau) = \frac{\int x_L(t)x_R(t+\tau)dt}{E_L E_R},\tag{1}$$

where x_L is the signal in the left ear, x_R is the signal in right ear, E_L is the root-mean-square energy of the signal in the

left ear, E_R is the root-mean-square energy of the signal in right ear, and τ is the ITD. If ρ is calculated over the entire stimulus duration, any long-term average ITD will only shift the position of the maximum value of the interaural cross-correlation function (which is also called the interaural coherence), and any long-term average ILD will not affect the maximum value of the interaural cross-correlation function because of the normalization by the energy in each ear. Therefore the interaural cross-correlation is mathematically related to the short-term ITD and ILD fluctuations that occur in the stimulus (e.g., Goupell, 2010).

Interaural decorrelation naturally occurs in realistic sound fields because of reflections in a room or the presence of multiple sound sources in a complex auditory environment. In normal-hearing (NH) listeners, interaural decorrelation is perceived as a diffuseness in the head (Whitmer *et al.*, 2012). The sensitivity of listeners to changes in interaural correlation appears to be closely related to the binaural masking level difference (BMLD) (Koehnke *et al.*, 1986; Goupell and Litovsky, 2014), which can be considered a simplification of the situation where listeners demonstrate improved speech understanding in noise for spatially separated talkers vs co-located talkers (e.g., Bronkhorst and Plomp, 1988; Culling *et al.*, 2004; Hawley *et al.*, 2004). In fact, binaural models that incorporate BMLD sensitivity can make extremely accurate speech-in-noise intelligibility predictions for a variety of spatial configurations, types of listening environments, and types of listeners (Lavandier and Culling, 2010; Culling *et al.*, 2012).

[a] Author to whom correspondence should be addressed. Electronic mail: goupell@umd.edu

Because binaural hearing provides the advantages of better sound localization and speech understanding in noise, it is not surprising that the clinical incidence of bilateral cochlear implants (CIs) has increased (Litovsky *et al.*, 2006; Litovsky *et al.*, 2009). However, benefits of binaural hearing are maximized when ITDs, ILDs, and interaural decorrelation cues are preserved and salient. At present bilateral CI users are faced with several barriers that prevent the availability of these binaural cues when using clinical speech processors.

CI processing strategies use vocoder-centric stimulus processing for stimulation strategies (Loizou, 2006). In such strategies, acoustic signals are bandpass filtered into contiguous channels. The envelope of each channel is extracted and is used to modulate the amplitude of electrical pulses. Thus typical CI processing discards the acoustic fine-structure information and presents only envelope information. Because NH listeners are very sensitive to ITDs in the fine structure at low frequencies (Brughera *et al.*, 2013) and they demonstrate a low-frequency dominance of ITD cues when localizing complex signals (Wightman and Kistler, 1992), the lack of low-frequency ITD information presented to bilateral CI users is a major limiting factor in achieving binaural hearing on par with NH listeners. Furthermore, bilateral CIs act as two independent monaural processors. Thus stimulation from the speech processors is not coordinated between the ears with regard to the timing, which is likely to distort or jitter ITDs in the sound envelope and in the timing between pulses. Therefore bilateral CI processors likely produce fluctuating ITDs.

In addition, it is likely that ILDs are distorted across the ears because of dissimilar loudness growth functions as well as independent microphone responses and automatic gain controls (van Hoesel, 2012). For instance, the amplitudes of the electrical pulses at each electrode range between threshold (T) and comfortable (C) stimulation in clinical current units (CUs). The T and C levels are normally set by a clinician for a standard clinical CI map. It is typical for loudness growth functions to be steep; thus amplitude compression is often applied to compensate for the steep loudness growth. The interaction of the steep loudness growth and compression likely does not produce equal loudness growth functions across frequencies and across the ears. This is potentially problematic as equal loudness across the ears and the perceived location of sounds are not directly verified in the mapping process (Goupell *et al.*, 2013). Thus the neural encoding of ILDs is not precisely controlled in bilateral CIs. Therefore bilateral CI processors likely produce fluctuating ILDs for any modulated acoustic signal.

The primary issue of concern in this study is the fact that distorted and fluctuating ITDs and ILDs create interaural decorrelation because the interaural correlation of modulated signals is not well controlled when transforming the physical acoustical signals to electrical signals. Hence it is unlikely that clinical processors provide binaural information that is optimal for horizontal-plane sound localization, detection of signals in noise, and improved speech understanding in noise.

One way to control some of the interaural distortions is to use bilaterally synchronized research processors instead of clinical processors. For example, studies where constant-amplitude pulse trains presented on single pairs of electrodes that are matched in pitch show that bilateral CI users are sensitive to ITDs and ILDs, particularly if the listeners were post-lingually deafened (Litovsky *et al.*, 2010). Several studies with synchronized processors have also shown that bilateral CI users demonstrate positive BMLDs. That is, thresholds are lower for conditions with NoSπ (diotic noise, tone out-of-phase) or NoSτ (diotic noise, tone with a non-zero ITD) compared to NoSo (diotic noise, tone in-phase and zero ITD) (Long *et al.*, 2006; Van Deun *et al.*, 2009; Lu *et al.*, 2010; Lu *et al.*, 2011; Van Deun *et al.*, 2011). Long *et al.* measured average BMLDs of 29 dB for 125-Hz transposed tones (i.e., stimulation pulses were dropped at regular intervals to produce small gaps in stimulation) using single pitch-matched pairs of electrodes in four CI listeners. The average BMLD was reduced to 9 dB when the sound processing included an amplitude compression algorithm as is typically applied in CI processing strategies. Lu *et al.* (2010) measured an average BMLD of 4.6 dB for tones masked by narrowband noises in five CI listeners. Van Deun *et al.* (2009) measured an average BMLD of 6.4 dB for 125-Hz transposed tones masked by narrowband noise in six children. Van Deun *et al.* (2011) measured average BMLDs of 12, 8, and 2 dB for NoS($\tau = 4$ ms), NoS($\tau = 2$ ms), and NoS($\tau = 0.7$ ms), respectively, for 125-Hz transposed tones in noise in three young adolescents and four adults. Therefore BMLD studies have been effective at demonstrating that bilateral CI users are sensitive to changes in interaural correlation encoded in the envelopes of signals.

The purpose of the present study was to improve our understanding of sensitivity to interaural envelope correlation changes in bilateral CI users as an indicator of the potential ability of these listeners to experience binaural unmasking of speech. The approach taken in this study was to measure sensitivity to interaural envelope correlation in BMLD and envelope correlation change tasks when stimulating single pitch-matched electrode pairs. ITD just noticeable differences (JNDs) were also measured using constant amplitude pulse trains as a measure of binaural sensitivity in the absence of envelope modulations. If the bilateral CI listeners are sensitive to binaural differences and the envelope correlation cues are salient, then positive BMLDs should be observed, and listeners should be able to detect changes in interaural envelope correlation.

II. EXPERIMENT I: BMLDS

A. Listeners and equipment

Eleven post-lingually deafened bilateral CI users participated in this experiment. Details regarding their hearing histories and etiologies are shown in Table I. All had Nucleus-type implants with 24 electrodes (Nucleus 24, Freedom, or N5). This type of CI has approximately 0.75-mm electrode spacing. The intra-cochlear electrodes are numbered such that the most apical electrode is 22 and the most basal is 1, and there are two extra-cochlear ground

M. J. Goupell and R. Y. Litovsky: Envelope correlation in cochlear implantees

TABLE I. Listener hearing histories, pitch-matched electrode pairs, and ITD JNDs. Thresholds marked "ND" were not determinable, meaning that a percent correct of 71.7% was not achieved at an ITD = 2000 μs.

Subject	Age	Etiology	Duration (yr)			Electrode pairs (L/R)	ITD threshold (μs)
			Deafness (L/R)	HA use (L/R)	CI use (L/R)		
IAJ	64	Unknown	36/43	46/53	14/7	19/21	460
						14/14	240
IAZ	76	Unknown	7/5	16/18	5/3	20/20	590
						13/14	409
IBD	80	Meniere's	29/29	8/30	12/12	20/22	209
						12/12	124
IBF	59	Hereditary	9/9	14/14	3/5	20/22	92
						12/13	38
IBJ	26	Unknown	3/3	16/17	1/2	21/20	ND
						12/13	ND
IBK	70	Noise exposure	9/9	8/8	7/1	18/22	162
						14/13	186
IBL	64	Unknown	20/20	0/47	10/5	12/12	ND
						4/4	614
IBM	56	Unknown	36/22	20/16	0.5/4	20/20	391
						12/12	174
IBN	63	Unknown	4/4	59/50	1/10	18/20	600
IBU	56	Unknown	5/5	16/16	4/4	12/12	384
IBV	69	Illness	20/54	26/11	9/2	20/19	ND

electrodes. These CIs have a range of 0 to 255 clinical CUs and the CUs produce logarithmically spaced changes of microamperes.

Bilaterally synchronized electrical pulses at single electrode pairs were presented directly to the listeners via L34 speech processors controlled by the Nucleus Implant Communicator (NIC, Cochlear Ltd.; Sydney, Australia). The processors were attached to a personal computer and controlled by custom software run in MATLAB (the Mathworks; Massachusetts).

B. Stimuli

1. Constant-amplitude stimuli

Stimuli were biphasic, monopolar electrical pulse trains presented to single pairs of pitch-matched electrodes. Each phase of the biphasic pulse had a 25-μs phase duration and a 8-μs gap between anodic and cathodic phases of the pulse. The pulse trains were 300 ms in duration, had no temporal windowing, and were presented at a rate of 100 pulses per second (pps). The constant-amplitude stimuli were used in the loudness mapping, bilateral pitch matching, and ITD discrimination measurements (see Secs. II C 1 and II C 2).

2. Noise stimuli

Dual-channel diotic narrowband analog Gaussian noises (No) started as the same stimuli that were used in Goupell (2012) and Goupell and Litovsky (2014). Stimuli were 500 ms in duration and temporally shaped by a Tukey window with a rise-fall time of 10 ms. The diotic masking noise had a 500-Hz center frequency. The bandwidth (BW) of the diotic masking noise was 10 or 50 Hz and was arithmetically centered on the center frequency.

The target stimuli consisted of a masking noise with a 500-Hz sine tone added either in-phase between the two ears (NoSo) or out-of-phase (NoSπ). The signal-to-noise ratio (SNR) of the level of the sine tone was varied relative to the overall level of the masking noise. The non-target stimuli were diotic noises (No). The stimuli were generated offline before the experiments. There were 25 different noise tokens for each condition and each value of SNR used in the experiments.

To convert the analog waveforms into electrical pulse trains, the Hilbert envelope was extracted for each channel. The envelope was then sampled at equal intervals corresponding to 1000 pps. The envelopes were compressed by a function similar to the one used by Long et al. (2006), which was an approximation of the compression used by Cochlear-type speech processors. This compression function is designed to transform the typically highly expansive loudness growth function of a CI listener to a loudness growth function that is more similar to a NH listener.[1] Therefore the electrical amplitudes (A_{El}) of the individual pulses were determined by the equation,

$$A_{El} = \begin{cases} \text{round}\left\{[1 - \exp(-5.09 E_{Ac})] \times (C - T) + T\right\}, & E_{Ac} \geq \text{Max}(E_{Ac}) - 30\,\text{dB} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where E_{Ac} is the instantaneous normalized analog envelope, C is the comfortable level in CUs, T is the threshold of hearing in CUs. To summarize, the analog amplitudes were compressed, quantized, and placed on a CU scale between T and C, and levels 30 dB below the average peak envelope amplitude were set to zero. Note that unlike Long *et al.* (2006) and Lu *et al.* (2010), the analog stimuli were not normalized to produce a peak amplitude that was unity for every token. Instead the analog stimuli were normalized such that the average peak amplitude over the 25 tokens corresponded to unity amplitude. This resulted in the flattening of the temporal peaks of some of the noise tokens. The reason for energy normalization was to better control the overall level of the stimuli. Due to the compression in the amplitude mapping function, it was not possible for electrical amplitudes to exceed a listener's C level. A portion of an example stimulus is shown in Fig. 1.

The signals were presented directly to listeners at bilaterally pitch-matched pairs of electrodes. The BMLD measurements were made at the electrode pair that had been shown to produce the best (lowest) ITD JND. The only exception was that listener IBK was tested with electrodes located in the middle of the electrode array because the ITD JNDs were very similar in the apex (162 μs) and middle (186 μs) (see Table I).

C. Procedure

The method of mapping, finding pitch-matched pairs of electrodes, and performing ITD JND measurements was similar to those described in Litovsky *et al.* (2012).

1. Loudness mapping

The T and C levels for each of the electrodes in both ears were used to determine loudness maps. Listeners reported the perceived loudness of a pulse train with the CUs varied incrementally, while ensuring comfortable levels at all times. After determining a loudness map, the C levels were compared across electrodes by sequentially playing the pulse trains with an interstimulus interval of 100 ms. Any electrodes that were perceived as softer or louder than the other electrodes had the current levels adjusted so that all of the electrodes had C levels that were equally loud. Three separate loudness maps were made for the following constant-amplitude pulse trains: 100 pps, 300 ms; 100 pps, 500 ms; and 1000 pps, 500 ms.

2. Bilateral pitch matching

A place-pitch magnitude estimation was performed using a method of constant stimuli. On each trial, listeners were presented with 100-pps, 300-ms constant-amplitude pulse trains at a single electrode in either the left or right ear. Only the even-numbered electrodes from 2 to 22 were tested. The amplitudes were the C levels from the loudness maps. Ten repetitions of the stimuli were presented for each condition in a randomized order. Therefore listeners were presented 220 stimuli (11 electrodes $\times$ 2 ears $\times$ 10 repetitions). Task familiarization was given to the listeners in an attempt to have them utilize the full range of the scale. Listeners rated the perceived place pitch of the stimulus on a scale from 1 to 100 and were given the option to repeat the stimulus before making their judgment. Listeners were also instructed to use the same scale for both ears. The results of the place-pitch magnitude estimation were used to pick approximately pitch-matched electrode pairs across the ears. The pitch-matched electrode pairs could only be estimated because of the great variability in the responses in the magnitude estimation task.
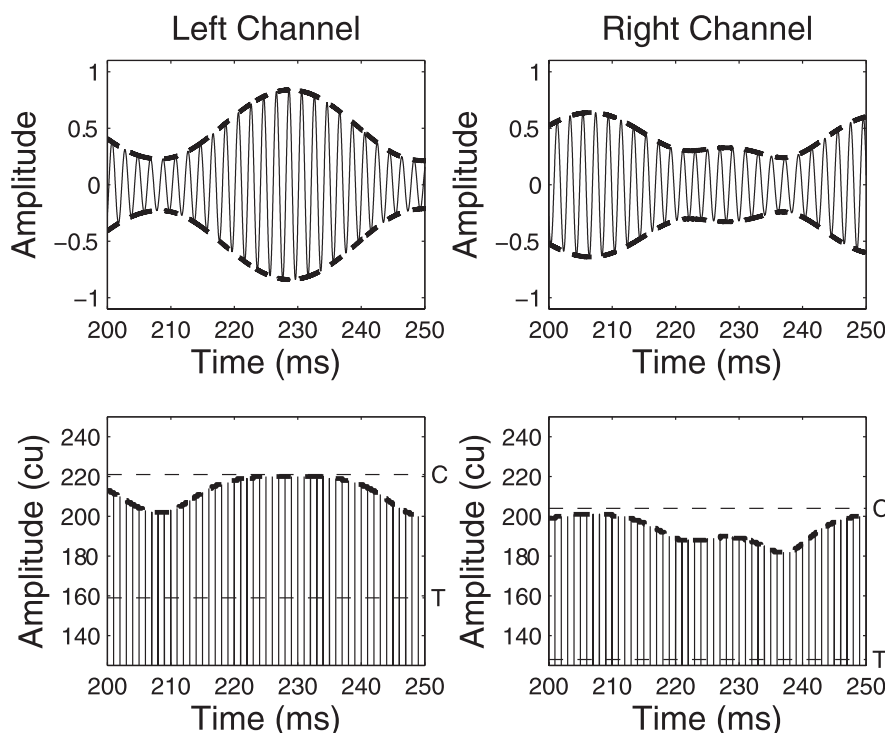

FIG. 1. A 100-ms portion of analog waveforms for the left and right channels (top row) and electrical pulse trains (bottom row) for a 50-Hz BW and 0-dB SNR NoSπ stimulus. The electrical pulse trains were processed using listener IBK's T and C levels for electrodes 18 (left) and 22 (right), which are shown by the horizontal dashed lines.

M. J. Goupell and R. Y. Litovsky: Envelope correlation in cochlear implantees

Next a direct left-right pitch comparison was performed. Taking the estimated pitch-matched electrode pairs, the electrode in the left or right ear was fixed (depending on the preference of the listener). The corresponding estimated pitch-matched electrode in the other ear, the adjacent electrodes (± 1 electrode from the estimated match), and the next adjacent electrodes (± 2 electrodes from the estimated match) were chosen for the direct pitch comparison. For listeners who had a poorer ability to discriminate place pitch, the range was doubled so that electrodes ± 2 and ± 4 from the estimated pitch-matched pair were tested. Listeners were presented 100-pps, 300-ms constant-amplitude pulse trains at C level in the left ear then in the right ear with a 300-ms interstimulus interval. The task of the listener was to indicate whether the stimulus in the right ear was "much lower," "lower," "the same," "higher," or "much higher" in pitch than the stimulus in the left ear. Listeners were given the option to repeat the stimulus to make their subjective judgment. At least 20 trials were presented for each combination of electrodes. Five places along the electrode array were tested (roughly apical, mid-apical, middle, mid-basal, and basal electrodes). The responses were converted to numerical scores where much lower was -2, lower was -1, the same was 0, higher was $+1$, and much higher was $+2$. The sum of the responses, μ, was calculated. The final pitch-matched electrode pair was the pair in which μ was closest to zero. If this method did not yield a definitive pitch-matched electrode pair (i.e., μ was equally close to zero for a number of electrode pairs or there was a non-monotonic change in μ over the range of electrodes), the combination closest in electrode number was chosen. When possible, we chose two pairs of electrodes for the following experiments. We chose one near the apical end of the electrode array and one near the middle because these regions typically contain vowel formants that are important for speech understanding. The pitch-matched electrode pairs for each listener are shown in Table I.

3. ITD sensitivity measurements at 100 pps (unmodulated pulse trains)

A method of constant stimuli was used to measure ITD JNDs at each pitch-matched electrode pair. Listeners were presented 100-pps, 300-ms constant-amplitude pulse trains. There were two intervals in each trial. One interval had a left-leading ITD, the other a right-leading ITD, and there was a 300-ms inter-interval duration. Because the pulse trains were unmodulated, the ITD was imparted on the pulse timing (i.e., fine structure). The task was to choose whether the stimulus had a left- or right-leading ITD for the second interval. The percentage of correct responses (PC) was calculated and psychometric functions with at least four points were made, each point consisting of at least 40 trials. Listeners were initially presented stimuli with ITDs of 100, 200, 400, and 800 μs. These values were adjusted depending on the overall sensitivity of the listener. The maximum ITD tested was 2000 μs. The JND was determined by finding PC = 70.7% using a maximum-likelihood fit (Wichmann and Hill, 2001). If listeners did not achieve threshold

performance for ITD = 2000 μs, the JND was classified as not determinable. No feedback was given after each trial, but a training period with feedback was initially provided until performance saturated. ITD JNDs are reported for each pitch-matched electrode pair in Table I.

4. BMLD measurements at 1000 pps (modulated pulse trains)

A four-interval, two-alternative forced-choice task was used. The listener initiated each trial, which consisted of four intervals grouped into sequential pairs. The inter-interval duration was 250 ms between stimuli in the first pair and in the second pair and was 500 ms between the two pairs. Three intervals contained non-target stimuli (No). The target stimulus (NoSo or NoSπ) was in either the second or third interval, which was chosen randomly. All four intervals had different stimulus tokens. No feedback was given after each trial, but a training period with feedback was initially provided until performance appeared to saturate. The NoSo conditions were tested first and then the NoSπ conditions. Ten listeners performed the 50-Hz BW conditions; six listeners performed the 10-Hz BW conditions.

Unlike Goupell (2012) and Goupell and Litovsky (2014) where an adaptive procedure was used to measure thresholds in NH listeners, a method of constant stimuli was used for the CI listeners to ensure a monotonically increasing psychometric function. Psychometric functions with at least three points were made, each point consisting of at least 40 trials. The SNRs tested typically had 6-dB spacing and the maximum SNR tested was $+12$ dB. Threshold was determined by finding PC = 70.7% using a maximum-likelihood fit. If listeners did not achieve threshold for $+12$-dB SNR, the threshold was classified as not determinable.[2]

D. Results and discussion

Nine of the 11 listeners were sensitive to ITDs for at least one pair of electrodes, which is shown in Table I.[3] Thresholds and BMLDs are shown in Fig. 2. Ten listeners had measurable thresholds for all of the conditions tested with the exception of listener IBV for the 10-Hz NoSo condition at $+12$-dB SNR. For this condition, the threshold was set to $+12$ dB for averages and statistical calculations. For the 10-Hz noises, the average NoSo and NoSπ thresholds were 3.2 and -7.9 dB, respectively, yielding an average BMLD of 11.1 ± 6.5 dB. For the 50-Hz noises, the average NoSo and NoSπ thresholds were -3.5 and -12.1 dB, respectively, yielding an average BMLD of 8.5 ± 4.2 dB. A two-way analysis of variance (ANOVA) with stimulus BW and configuration (NoSo and NoSπ) as the factors yielded a significant effect of BW [$F_{(1,32)} = 6.2$; $p = 0.019$] and configuration [$F_{(1,32)} = 20.2$; $p < 0.0001$], but the interaction was not significant ($p > 0.05$).

The average BMLD for tones in noise measured here were found to be larger than the average BMLDs for tones in noise measured by Lu et al. (2010) and were more similar to the BMLDs for transposed tones measured by Long et al. (2006) and Van Deun et al. (2011). Although our stimuli are most like those in Lu et al. (there were no regular gaps in
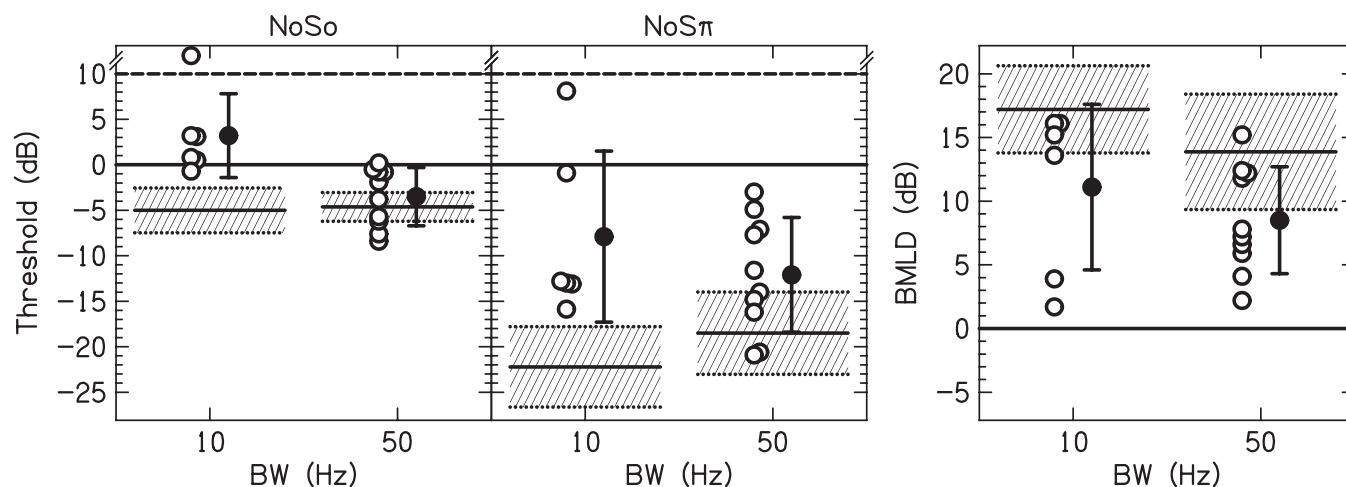
FIG. 2. Individual (open) and average (closed) thresholds and BMLDs. Error bars represent ±1 standard deviation. The mean threshold (solid line) and ±1 standard deviation (shaded area) for NH listeners presented a CI simulation are also shown.

stimulation as occur for transposed tones), several factors might explain the discrepancy between our data and those of Lu *et al*. First, we provided feedback and training prior to collecting data, whereas Lu *et al*. did not provide training. Second, we measured BMLDs at only one place on the electrode array, generally where the best ITD JND had been measured. Lu *et al*. tested at several places across the electrode array, which probably included electrodes pairs with both good and poor binaural sensitivity for an individual listener. If there is a correlation between NoSπ and ITD sensitivity, this could explain part of the discrepancy. Such a correlation was found in our data, as can be seen in Fig. 3(A). Omitting the one listener who did not have a determinable ITD JND, a linear regression between the 50-Hz NoSπ thresholds and ITD JNDs accounted for 67% ($p = 0.007$) of the variance in the data. A similar significant relationship between NoSπ thresholds and ITD JNDs was found by Long *et al*. (2007). Third, we measured psychometric functions, whereas Lu *et al*. performed an adaptive tracking procedure to determine threshold. If some of the psychometric functions were non-monotonic, those thresholds measured in an adaptive procedure may be unreliable. We did measure one non-monotonic psychometric function for listener IAZ for the 10-Hz NoSπ condition. To understand how a non-monotonic psychometric function could

occur for this condition, we performed an in-depth analysis of the stimuli used in our experiment.

E. Stimulus analysis

The NoSo and NoSπ detection tasks could potentially be performed based on a number of detection cues. These cues may change as a function of SNR (e.g., a transition from monaural to binaural cues for NoSπ detection as SNR decreases). In addition, because of the differences between analog and electrical stimulation, there is no guarantee that NH and CI listeners are being presented similar monaural and binaural detection cues. Hence an analysis of the potential detection cues present in the electrical stimuli was performed to provide insight into the CI listener performance.

Figure 4 shows the stimulus analysis for the signals used in this experiment. The analyses of the analog signals (including both fine structure and envelope changes) and electric signals (including envelope changes, but the pulse timing is synchronized so there are no fine structure changes) are on the left and right, respectively. The separate columns are for the 10- and 50-Hz BW stimuli. Note the different scales between the analog and electric stimuli. Four metrics were computed, and the average of each metric over 25 stimulus tokens was calculated for No (dashed line), NoSo (open symbols), and NoSπ (closed symbols). For the electric signals, the stimuli
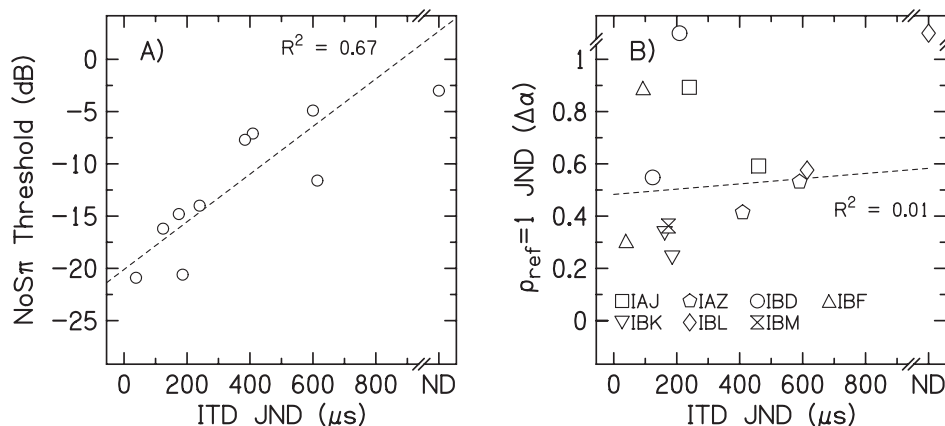

FIG. 3. The relationship between NoSπ thresholds and ITD JNDs (A) and $\rho_{ref} = 1$ and ITD JNDs (B) for 50-Hz BW stimuli. Linear regressions are represented by the dashed lines, which did not include the not determinable JNDs. In (B), different symbols represent JNDs for individual listeners.
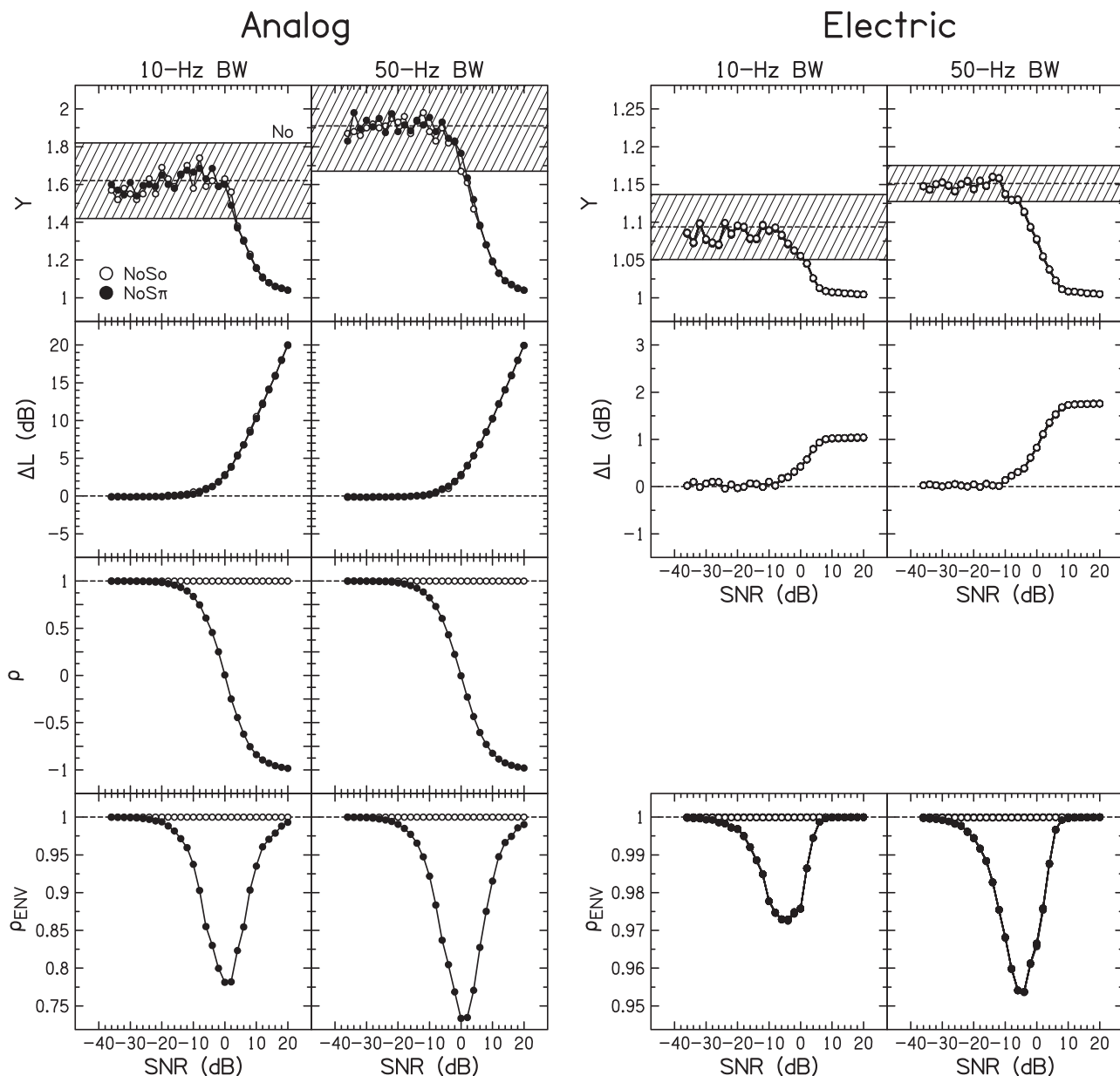
Analog

10−Hz BW   50−Hz BW

Electric

10−Hz BW   50−Hz BW

FIG. 4. Calculation of the normalized kurtosis (Y), change in level (ΔL), waveform interaural correlation (ρ), and interaural envelope correlation (ρ_{ENV}) as a function of SNR for the analog stimuli and after transformation to electrical CUs. Open symbols show the average value for NoSo stimuli; closed symbols show the average value for NoSπ stimuli. The shaded region in the panels depicting Y shows the mean and ± 1 standard deviation for No stimuli. Note that the electrical stimuli demonstrate smaller ranges of each stimulus metric compared to the analog panels. Data points for each CI listener using their particular DR are plotted in the electric panels but are not visible because DR only slightly changed the calculation of each metric. Calculations for the NoSπ stimuli were omitted from the electric Y and ΔL panels to improve visibility.

were passed through the compression algorithm in Eq. (2) for all of the listeners' individual T and C levels. Data points for the metrics of the electric stimuli are nearly identical across individual listeners, hence not easily visible in Fig. 4.

1. Monaural stimulus metrics

The first metric was the normalized kurtosis,

$$Y = \frac{\langle E^4 \rangle}{\langle E^2 \rangle^2},\tag{3}$$

where E is the envelope and $\langle x \rangle$ denotes the expectation of x. For a tone with a perfectly flat envelope, the value of $Y = 1$.

If a listener used changes in the envelope fluctuations (i.e., envelope roughness) as a detection cue, changes in Y would be related to listener performance.

In the top row of Fig. 4, the general shapes of the curves are the same between the analog and electric stimuli. For positive SNRs, the envelope of the composite tone in noise was more similar to a sine tone, which is perfectly flat ($Y = 1$). As the SNR decreased, the envelope of the composite in noise became more like the envelope of the noise, thus trending toward the average Y for the No stimuli. The noises had Gaussian-distributed instantaneous amplitudes, and the average envelope fluctuation rate of the composite sound was approximately 64% of the BW of the noise (Rice, 1954). Because this is a monaural metric, Y behaved

similarly the for the NoSo and NoSπ stimuli. Thus NoSπ curves were not plotted for the electric stimuli.

The average Y for the 10- and 50-Hz analog NoSo and NoSπ stimuli was less than the No range at +4- and +2-dB SNR, respectively. The average Y for the 10- and 50-Hz electric NoSo and NoSπ stimuli was less than the No range at +2- and −4-dB SNR, respectively. Therefore Y might be usable as a detection cue at slightly smaller SNRs for CI than NH listeners. The average Y continued to decrease for increasing SNR for both the analog and electric stimuli, and if either the NH or the CI listeners attempted to utilize Y to perform the detection task, that cue would not become more salient with increasing SNR. In fact for the CI listeners, the value of Y remained unchanged at a value of approximately 1 for SNRs greater than +6-dB SNR.

The second monaural stimulus metric was the change in the level in decibels from NoSo or NoSπ compared to No (ΔL). These calculations were performed using the analog amplitude or electric CUs. For stimuli with no change, $\Delta L = 0$ dB. If a listener used changes in the intensity as a detection cue, ΔL of the stimuli would be related to the performance.

In the second row of Fig. 4, the average ΔL became larger than zero in the range of −10- to 0-dB SNR. There is one noticeable difference in the functional forms of ΔL vs SNR between the analog and electric stimuli. The average ΔL continually increased for increasing SNR for the analog stimuli. However, the average ΔL saturated at 1 or 2 dB above about +6-dB SNR for the electrical stimuli, which can be explained by the signal compression [see Eq. (2)]. Therefore if CI listeners were not able to detect ΔL at +6-dB SNR, they would not be able to use that cue for larger SNRs. This may explain how one of the NoSo thresholds was not determinable for 10-Hz BW. Also note that the increase in ΔL as SNR increased is a result of the increase in the amplitude of pulses encoding valleys of the noise envelope modulations. It is not an increase in the amplitude of pulses encoding the peaks of the noise envelope modulations; the compression in Eq. (2) limited the maximum amplitude to C level (see Fig. 1). Therefore to understand the perceptual relevance of the change in ΔL, one must consider the change in loudness produced by the increase of the individual pulses in the valleys of the temporal waveform.

Following testing, listeners were asked to provide anecdotal feedback on whether they used envelope roughness (i.e., changes in monaural fluctuations), loudness, or both cues to perform the NoSo detection task. All of the CI listeners reported that they primarily used roughness cues. This is in contrast to NH listeners who, in previous studies, reported that they primarily used loudness cues (Goupell, 2012; Goupell and Litovsky, 2014).

2. Binaural stimulus metrics

The third stimulus metric was the normalized cross-correlation of the waveform [ρ, Eq. (1)], which has a value of $\rho = 1$ for interaurally correlated signals, $\rho = 0$ for interaurally uncorrelated signals, and $\rho = -1$ for interaurally anti-correlated (i.e., out of phase) signals. Because $\rho = 1$ for

the carrier pulse train of the electrical signals, it was omitted from the right columns of the third row of Fig. 4. The value of ρ remained 1 for the NoSo stimuli. The value of ρ changed from near 1 for negative SNRs, to near 0 for 0-dB SNR, to near −1 for positive SNRs for NoSπ stimuli.

The fourth stimulus metric was the normalized envelope cross-correlation (ρ_{ENV}), which cannot achieve values below zero because the envelope by definition has only positive values. The envelope was calculated using the Hilbert transform. In the fourth row of Fig. 4, the average ρ_{ENV} is plotted for both the analog and electric stimuli. By definition, the NoSo stimuli have $\rho_{\text{ENV}} = 1$. If a listener used changes in the interaural envelope correlation as a detection cue, changes in ρ_{ENV} would be related to listener performance.

For the NoSπ analog stimuli, there was a minimum in the ρ_{ENV} function at 0-dB SNR. The curve for the smallest and largest SNRs approached $\rho_{\text{ENV}} = 1$ because the envelopes of the stimuli in the right and left ears become increasingly similar. For the electric stimuli, the NoSπ stimuli demonstrated a minimum in the ρ_{ENV} function closer to −5-dB SNR. Note that the values in the ρ_{ENV} function were much nearer to 1 for the electric stimuli than the analog stimuli. This occurred because of the compression in the signal processing in Eq. (2). Quantization of the instantaneous amplitudes, before or after compression, produced very little change in the values of ρ_{ENV}. As mentioned in the preceding text, there was a negligible effect of dynamic range (DR) in the values of ρ_{ENV}.

3. Stimulus analysis summary

To perform the NoSo task, listeners needed to discriminate NoSo from No based on the monaural cues that were present. A lower NoSπ threshold compared to the NoSo threshold must be based on additional binaural cues produced by a difference in ρ_{ENV}. The differences between the analog and electrical stimuli that resulted from the limited DR and compression function used in CI listeners has the potential to affect the salience of the detection cues. To summarize the analysis, for the monaural cues, the compression function in Eq. (2) changed how the listeners might perform the task, particularly compared to NH listeners who do not experience the same compression. The compression limited the ability CI listeners to discriminate No from NoSo or NoSπ by using ΔL (i.e., an intensity cue) for large, positive SNRs. Discriminating No from NoSo or NoSπ could have been done using Y (i.e., envelope or roughness cues) at large, positive SNRs; this is consistent with the report of the CI listeners' subjective report regarding which cue they used to perform the task. Assuming NH listeners utilized intensity cues to perform NoSo and NoSπ detection at relatively larger SNRs, CI listeners seems to have used a different detection cue by using the envelope roughness.

The envelope cues occur at a slower rate for the 10-Hz BW stimuli compared to the 50-Hz BW stimuli. This slower rate might make them less salient for CI listeners; this may explain the higher NoSo thresholds for the 10-Hz BW compared to the 50-Hz BW. It also may explain why one NoSo threshold was not determinable at the 10-Hz BW. Note that

there was no effect of stimulus BW on NoSo thresholds in NH listeners presented a CI simulation (Goupell and Litovsky, 2014); this again suggests that NH listeners utilized the intensity cue and disregarded the envelope cues for NoSo and NoSπ detection at relatively larger SNRs.

For the potential binaural detection cues, the function is non-monotonic, which also occurs for analog signals (Bernstein and Trahiotis, 1996). At large, positive SNRs, there is no difference in ρ_{ENV} for the NoSo and NoSπ stimuli, as expected. Therefore if CI listeners cannot use the monaural cues and the binaural cues are non-existent for large, positive SNRs, this could explain how non-monotonic NoSπ psychometric functions could be produced for CI listeners but not NH listeners.[4]

It should be noted that these calculations were simply conducted on the analog and electric stimuli to understand the physical properties of the stimuli before neural encoding. The acoustic-to-electric compression function in sound processing strategies simulated in Eq. (2) ensures that a large acoustic DR (typically 40 dB in modern processors, which encodes a range of levels that are most important for speech understanding) can be accommodated by the relatively small electric DR (typically 5–10 dB, which is limited by the expansive electrical loudness growth functions). Accordingly, a 1-dB change in electric stimulation for CI listeners will likely produce much larger perceptual effect than a 1-dB change for acoustic stimulation for NH listeners. It would be incorrect to expect a 1-dB detection cue should be equally salient in both groups of listeners.

III. EXPERIMENT 2: ENVELOPE CORRELATION CHANGE DISCRIMINATION JNDS

The results of experiment 1 demonstrated that bilateral CI listeners are sensitive to changes in the interaural envelope correlation because the NoSπ thresholds were better than the NoSo thresholds. However, there are potentially both monaural and binaural cues available to perform the NoSπ detection task. In experiment 2, we used noises that were varied in interaural correlation without the addition of a sine tone, effectively reducing potential monaural detection cues such as envelope roughness or loudness. For the NoSπ stimuli, the interaural envelope correlation is a non-monotonic function (see Fig. 4). Therefore another reason for performing this experiment was to enable us to more precisely control the interaural correlation and utilize stimuli that had a monotonic interaural envelope correlation function. We investigated the effects of masker BW (i.e., average envelope modulation rate), pulse rate, and detection of changes from different reference correlations.

A. Method

Seven listeners from experiment 1 participated in this experiment (IAJ, IAZ, IBD, IBF, IBK, IBL, and IBM).[5] Tests were performed at one or two pitch-matched electrode pairs (see Table I), although not all listeners were tested in each condition because of time constraints.

The analog stimuli used to generate the electrical signals had a 500-Hz center frequency and a 10-, 50-, or 400-Hz

BW. The duration of the stimuli was 500 ms. They were temporally shaped by a Tukey window with a 10-ms rise-fall time. The CI processing and compression function was the same as in experiment 1. The target stimuli for the envelope correlation change discrimination tasks had a varied value of the cross-correlation (ρ), while the reference stimuli had one of two reference correlations ($\rho_{ref} = 1$ or 0). Various levels of decorrelation were achieved by orthogonalization of the noises and precisely controlling the target correlation on a scale of $\alpha = \sqrt{1 - \rho^2}$ with 0.05-α steps (Culling et al., 2001; Goupell, 2010). Note that for these stimuli, α (concomitantly ρ) is highly controlled, and this will produce systematic changes in ρ_{ENV}. An analysis of the two variables showed that there is a monotonic decrease in ρ_{ENV} as α increases as well as a fairly linear relationship between the two variables for values of $\alpha \geq 0.4$.

The forced-choice procedure was the same as the one used in experiment 1. The non-target stimuli had a correlation of $\rho_{ref} = 1$ or 0, depending on the condition. The target stimuli had a change in correlation from the reference. At least three-point psychometric functions with at least 40 trials per point were measured. In conditions where PC = 70.7% was not achieved, the JND was set to a value of $\Delta\alpha = 1.1$, corresponding to the value used in previous studies that performed similar measurements in NH listeners (Goupell, 2012; Goupell and Litovsky, 2014). In all analyses, we did not take into account the cochlear place of stimulation because of the large variability across place within single CI listeners, which often results in a lack of significant effect of place (e.g., Litovsky et al., 2010; Lu et al., 2010).

B. Results

The results of the varied masker BW are shown in Fig. 5 for all seven listeners. All conditions are for a 1000-pps pulse rate and $\rho_{ref} = 1$. JNDs for different pairs of electrodes are shown by open (apical or basal pair) or closed (middle pair) symbols. In a one-way ANOVA with BW as the factor, there was not a significant effect of BW ($p > 0.05$). There
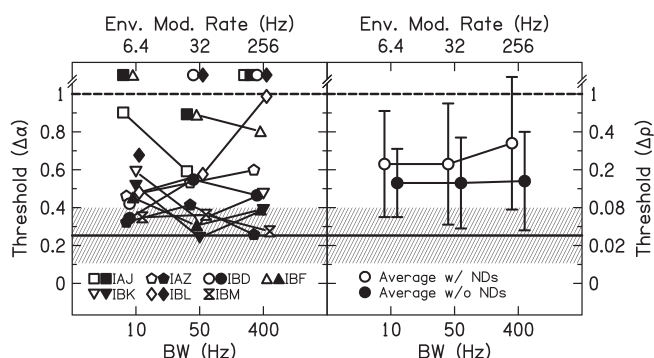

FIG. 5. Envelope correlation change JNDs for $\rho_{ref} = 1$ non-targets and 1000-pps stimuli as a function of stimulus BW. Individual (left panel) and average (right panel) JNDs are shown. In the left panel, open and closed symbols show measurements made at different places along the electrode array. In the right panel, averages included (open symbols) or excluded (closed symbols) not determinable JNDs and error bars represent ±1 standard deviation. The mean JND (solid line) and ±1 standard deviation (shaded area) for NH listeners presented a CI simulation are also shown.
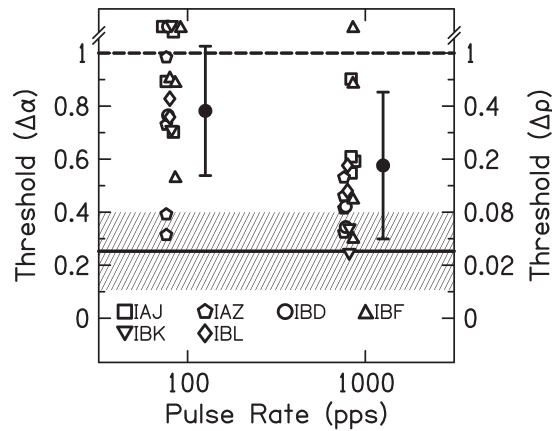
FIG. 6. Envelope correlation change JNDs for $\rho_{ref} = 1$ non-targets as a function of pulse rate. Individual (open) and average (closed) JNDs are shown. Error bars represent ± 1 standard deviation. The mean JND (solid line) and ± 1 standard deviation (shaded area) for NH listeners presented a CI simulation are also shown.

was also no effect if the not determinable JNDs were removed from the calculation.

The results of the varied pulse rate are shown in Fig. 6. All conditions are for $\rho_{ref} = 1$. The stimuli had a BW of 10 or 50 Hz, but because there was no effect of BW in Fig. 5, both BWs are plotted together. Six listeners participated with one or two electrode pairs and one or two BWs (IAJ: 2 pairs/2 BWs, IAZ: 2 pairs/2 BWs, IBD: 1 pairs/2 BWs, IBF: 2 pairs/2 BWs, IBK: 2 pairs/1 BW, and IBL: 1 pair/2 BWs). Some of the data from Fig. 5 were included in this analysis. In a two-way ANOVA with pulse rate and BW as the factors, there was a significant effect of pulse rate [$F(1,35) = 5.53$, $p = 0.025$] where increasing pulse rate produced lower JNDs. The effect of BW and the interaction of rate × BW were not significant ($p > 0.05$).

The results of the varied reference correlation are shown in Fig. 7. All conditions are for a 1000-pps pulse rate. The 10-Hz BW conditions consisted of six measurements (two
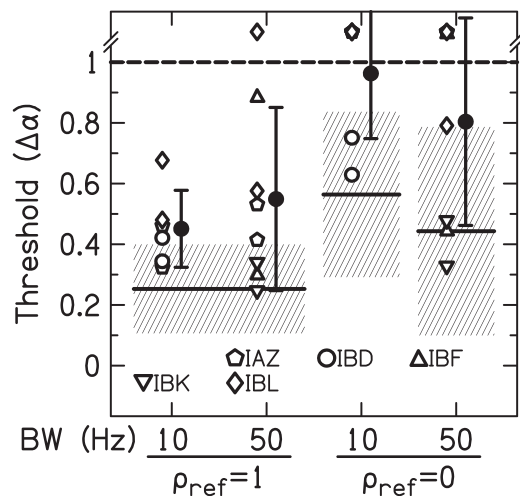

FIG. 7. Envelope correlation change JNDs for 1000-pps stimuli as a function of reference correlation. Individual (open) and average (closed) JNDs are shown. Error bars represent ± 1 standard deviation. The mean JND (solid line) and ± 1 standard deviation (shaded area) for NH listeners presented a CI simulation are also shown.

pairs IAZ, two pairs IBD, and two pairs IBL); the 50-Hz BW conditions consisted of eight measurements (two pairs IAZ, two pairs IBF, two pairs IBK, and two pairs IBL). Some of the data from Figs. 5 and 6 were included in this analysis. In a two-way ANOVA with reference and BW as the factors, there was a significant effect of reference [$F(1,27) = 13.1$, $p = 0.001$] where $\rho_{ref} = 1$ JNDs were lower than $\rho_{ref} = 0$ JNDs. The effect of BW and the interaction of reference × BW were not significant ($p > 0.05$).

Table II shows the NoSπ thresholds and $\rho_{ref} = 1$ JNDs for the six listeners who performed these measurements at the same pair of electrodes and the same BWs. Their thresholds and JNDs were converted to analog values of ρ_{ENV} (see Fig. 4) to determine if the listeners might have performed the two tasks in a similar way. When averaged over all listeners, experiment types, and BWs, the change in correlation necessary to achieve threshold was $\Delta\rho_{ENV} = 0.04$. For the 12 comparisons across experiment types in Table II, only three differences produced $|\Delta\rho_{ENV}| > 0.04$. Therefore there is a good correspondence between the values of ρ_{ENV} for the NoSπ thresholds and $\rho_{ref} = 1$ JNDs; this suggests that the listeners were mostly relying on the interaural envelope correlation to perform both tasks.

Figure 2(B) shows the 50-Hz BW, 1000-pps, $\rho_{ref} = 1$ JNDs plotted against ITD JNDs for the seven listeners in this experiment. Omitting the not determinable JNDs, the correlation between these two measurements was not significant ($R^2 = 0.01$, $p = 0.73$). In addition, the $\rho_{ref} = 1$ JNDs were converted from α to ρ and ρ_{ENV}. The correlations between the $\rho_{ref} = 1$ JNDs using the other metrics and the ITD JNDs were not different than the correlations with α.

C. Discussion

This experiment investigated the sensitivity of bilateral CI listeners to changes in interaural envelope correlation in the absence of monaural detection cues that are introduced when measuring NoSπ thresholds. Similar to experiment 1, bilateral CI listeners were mostly sensitive to changes in interaural envelope correlation, and there was large inter-individual variability. Comparisons between experiment 1 and 2 showed good correspondence between the values of ρ_{ENV} for the NoSπ thresholds and $\rho_{ref} = 1$ JNDs (Table II). Therefore the CI listeners were likely using the interaural envelope correlation to perform both tasks, as occurs for NH listeners (Goupell and Litovsky, 2014). In contrast to experiment 1 where there was always a positive BMLD, there were many conditions in this experiment where individual listeners did not demonstrate sensitivity to interaural envelope correlation. In addition, the thresholds were on average lower for the NoSπ detection task compared to the correlation change discrimination task ($\Delta\rho_{ENV} = 0.03$ for the 10-Hz BW and $\Delta\rho_{ENV} = 0.01$ for the 50-Hz BW). This may be an indication that the presence of monaural cues for the NoSπ detection task provided a slight advantage for some of the CI listeners. Alternatively, this discrepancy could have occurred because measurements were made at only the place of stimulation along the cochlea that had the highest binaural sensitivity (experiment 1), compared with measurements at two

TABLE II. Comparison of ρ_{ENV} for NoSπ thresholds from experiment 1 and $\rho_{ref} = 1$ correlation change JNDs from experiment 2. The difference in ρ_{ENV} (denoted $\Delta\rho_{ENV}$) for the NoSπ thresholds and the correlation change JNDs is reported.

	10-Hz BW					50-Hz BW				
Listener	NoSπ threshold (dB)	NoSπ threshold (ρ_{ENV})	Correlation change JND ($\Delta\alpha$)	Correlation change JND (ρ_{ENV})	$\Delta\rho_{ENV}$	NoSπ threshold (dB)	NoSπ threshold (ρ_{ENV})	Correlation change JND ($\Delta\alpha$)	Correlation Change JND (ρ_{ENV})	$\Delta\rho_{ENV}$
IAJ	−13.0	0.97	1.1	0.83	0.14	−14.0	0.97	0.89	0.83	0.13
IAZ	−12.8	0.97	0.32	0.98	−0.01	−7.1	0.86	0.41	0.98	−0.12
IBD	−15.9	0.98	0.34	0.97	0.01	−16.2	0.98	0.55	0.94	0.04
IBF	−13.1	0.97	0.45	0.96	0.00	−20.9	0.99	0.31	0.98	0.01
IBK	−14.1	0.97	0.52	0.94	0.03	−20.6	0.99	0.24	0.99	0.01
IBM	−17.4	0.98	0.35	0.97	0.01	−14.8	0.97	0.36	0.97	0.00
Average	−14.4	0.97	0.51	0.94	0.03	−15.6	0.96	0.46	0.95	0.01

places, including some cases in which sensitivity was poorer at one of the places (experiment 2). However, this idea was not supported by the lack of correlation found between ITD and $\rho_{ref} = 1$ JNDs shown in Fig. 2(B).

On average, there was no effect of BW on $\rho_{ref} = 1$ JNDs (Fig. 5). Using a 100-pps carrier to encode the envelope from the 50-Hz BW noises produced higher $\rho_{ref} = 1$ JNDs than using 1000-pps carrier (Fig. 6). Two factors that may be related to the higher 100-pps JNDs are the DR and sampling of the envelope. The DR increases as the pulse rate increases, which may decrease JNDs at higher rates. For a 50-Hz BW noise, the average envelope modulation rate is approximately 32 Hz. While 100-pps carrier sampling of the signal satisfies the Nyquist criterion of using a sampling frequency of at least two times the highest frequency in the signal, it may be that four times the highest frequency is necessary for encoding of electrical signals (McKay *et al.*, 1994).

In experiment 2, uncorrelated reference ($\rho_{ref} = 0$) JNDs were also measured. Average $\rho_{ref} = 1$ JNDs were lower than $\rho_{ref} = 0$ JNDs (Fig. 6), which also occurs for NH listeners (Goupell, 2012; Goupell and Litovsky, 2014). In contrast to NH listeners, many of the $\rho_{ref} = 0$ JNDs were not determinable, which may be related to the overall poorer sensitivity to changes in correlation when only envelope cues are available.

IV. GENERAL DISCUSSION

A. Summary of experiments

These experiments were aimed at measuring the sensitivity of CI listeners to changes in interaural envelope correlation, which may ultimately be related to speech understanding in noise. Results showed that most bilateral CI listeners are sensitive to changes in interaural envelope correlation for both NoSπ stimuli and those generated with specified values of interaural correlation. The BW of the stimuli was a significant factor in experiment 1 (Fig. 2) but not in experiment 2 (Fig. 5). The results of experiment 2 are consistent with the results of Lu *et al.* (2010); however, both experiment 2 in this study and the experiment in the Lu *et al.* study may have been limited by the variability in the measurements.

There may be some evidence that DR is important for properly representing interaural envelope correlation in some listeners; this is consistent with the results of experiment 2. We found that 1000-pps stimuli produced lower envelope correlation change JNDs than 100-pps stimuli (Fig. 6), which may be a result of better representation (less quantization) of the instantaneous level for the higher stimulation rates. Other explanations for this difference include the loudness growth functions changing as the rate changes, inadequate sampling of the envelope at 100 pps but not at 1000 pps, and increased importance of the synchronized carrier pulses at 100 pps but not at 1000 pps. The fact that the high-rate pulse trains produced lower JNDs than the low-rate pulse trains is also in the opposite direction of the rate limitations shown in CI users for ITD discrimination of constant-amplitude pulse trains (Laback and Majdak, 2008; van Hoesel *et al.*, 2009). Note that these rate limitations are reduced or eliminated with envelope modulations (van Hoesel *et al.*, 2009; Noel and Eddington, 2013). Therefore it seems that the ITD-based rate limitation observed in bilateral CI listeners is solely a temporal processing phenomenon and does not affect the ILD-based encoding of interaural decorrelation.

When the bilateral CI listeners were tested using uncorrelated reference intervals, performance became very poor. In Fig. 7, almost all of the JNDs for $\rho_{ref} = 1$ (13/14) were measurable; however, less than half of the JNDs for $\rho_{ref} = 0$ (6/14) were measurable. Such poor performance from the CI listeners has implications for understanding speech in realistic environments, which will be discussed in Sec. IV C.

B. Comparisons to NH performance

It can be instructive to compare the results from the bilateral CI users to those obtained from NH listeners who are tested with CI simulations. In this case, the simulations were acoustic pulse trains that were designed to be similar to the electric pulse trains. The acoustic pulse trains encoded the temporal envelopes extracted from the narrowband noises. The pulses were synchronized across the ears. The pulses had a 2000-Hz bandwidth to simulate the large spread of current from monopolar stimulation. The pulses had a carrier with a 4000-Hz center frequency to avoid phase locking to the carrier. Therefore the pulses should represent the

modulated noise envelopes when the pulses were presented in a train. Acoustic pulse trains were presented at 65 dB-A, which produced a comfortable loudness similar to the comfortable-loudness pulse trains presented to the CI listeners. No compression was applied to the acoustic pulse trains because the compression described in Eq. (2) was intended to replicate loudness growth that occurs in normal hearing; adding compression would undo the intention of the compression applied in the CI stimulus processing. In summary, the pulse trains presented to the CI and NH listeners were made to present relatively similar envelope-based cues. However, note that many aspects of electrical stimulation cannot be mimicked acoustically and caution should be taken in such a comparison.

The acoustic pulse train data mostly come from other studies (Goupell, 2012; Goupell and Litovsky, 2014). The exception was the data for the 50-Hz BW NoSo and NoSπ conditions, which were collected for this study. The acoustic pulse train data were averaged over stimulation rate because there was no effect of pulse rate for these conditions (Goupell, 2012). The average and standard deviation of the NH listeners data are plotted in Figs. 2, 5, 6, and 7.

While some CI listeners presented electric pulse trains performed as well as the NH listeners presented acoustic pulse trains, performance for the CI listeners is on average worse than performance for the NH listeners. One signal-based explanation for this difference is that while the analog signals with a specific interaural envelope correlation were transformed into modulated electrical pulse trains with a *controlled* interaural envelope correlation, the neural representation of the interaural envelope correlation for those pulse trains was markedly *uncontrolled*. In other words, a correlated non-target stimulus (No or $\rho_{ref} = 1$) in the discrimination task was in fact decorrelated at the electrode-neural interface or some higher neural center. This internal form of interaural decorrelation could occur from the mapping of acoustic amplitude to the electric domain because it involves a non-individualized amplitude-to-current mapping function [see Eq. (2)]. For this non-individualized approach, the amount of envelope decorrelation was likely different across listeners, which may in part explain the large inter-listener variability seen in our CI population (see Figs. 2, 5, 6, and 7). NH listeners are extremely sensitive to changes in correlation from a perfectly correlated reference but much worse at detecting changes from a decorrelated references (Gabriel and Colburn, 1981; Koehnke et al., 1986; Culling et al., 2001; Goupell, 2012; Goupell and Litovsky, 2014). Thus a parsimonious explanation for our data is that some of our CI listeners are not necessarily less sensitive than NH listeners in this task, but they are demonstrating a typical finding, which is reduced sensitivity to changes in envelope correlation because of the internal decorrelation in the neural representation of the signals.

It is likely that the interaural envelope correlation of our stimuli was not well controlled because of imprecise envelope modulation encoding across the ears. The method of mapping we used, similar to clinical mapping procedures, involved measuring T and C levels without determination of underlying loudness growth functions, which may need to be measured to avoid spurious lateralizations and thus change the interaural correlation from the desired value (Goupell et al., 2013). The specific T and C levels, and perhaps the DRs between the ears, are often different, which occurs because of different neural survivals or physical placements of the CIs in the cochleae.

On the other hand, variability in the CI listener performance is common (van Hoesel et al., 2009; Lu et al., 2010; Litovsky et al., 2012), so not all of the variability should be attributed to the signal-based explanation of the data. There could be other non-signal-based factors that contribute to the relatively poor performance in the CI listeners tested with electrical pulse trains compared to the NH listeners tested with acoustical pulse trains. For example, the acoustical pulse trains might not be simulating the necessary features of electrical stimulation. For example, the CI simulation might have been less than optimal because there was a lack of low-frequency masking noise. Such noise is necessary to mask low-frequency distortion products that NH listeners can use to perform binaural tasks. In other words, NH listeners' thresholds and JNDs may have been artificially low. However, as argued in Goupell (2012), low-frequency masking noise might not be necessary in this simulation because the acoustic stimuli consisted of synchronized acoustic pulses that only represent the amplitude differences between the ears. Because only time-varying ILDs are being presented, there would be no obvious advantage for NH listeners to use the ILDs of the low-level, low-frequency distortion products over the high-level high-frequency signal because ILD sensitivity is mostly independent of frequency in NH listeners using headphones when presented non-vocoded stimuli (Yost and Dye, 1988). Hence we doubt that this is a major factor in the difference in the CI and NH binaural sensitivity.

Another explanation for the CI-NH discrepancy is that some listeners reported that the sound images were not perfectly intercranially centered. Studies using constant-amplitude pulse trains to measure ITD and ILD sensitivity often center the stimuli to maximize binaural performance (Litovsky et al., 2012) because centered images show the best binaural discrimination performance in NH listeners (Yost and Dye, 1988; Koehnke et al., 1995). Therefore the uncentered images may have decreased performance in the CI listeners, but not in the NH listeners.

Pitch-matched pairs of electrodes were tested in these experiments, which often maximize sensitivity to binaural cues in CI listeners (Long et al., 2003; Poon et al., 2009; Kan et al., 2013). Thus interaural place-of-stimulation mismatch may have also reduced binaural sensitivity in the CI listeners if the pitch matches did not fully compensate for interaural frequency mismatch.

Neural deprivation and neural degeneration could have also been a contributing factor (Bierer and Faulkner, 2010). One consequence of neural degeneration is decreased binaural sensitivity due to a fewer number of neurons encoding the signals. If neural degeneration is the main cause for the CI listener performance and not the uncontrolled envelope modulation encoding, one might expect a strong correlation between NoSπ thresholds using modulated pulse trains and

M. J. Goupell and R. Y. Litovsky: Envelope correlation in cochlear implantees

ITD JNDs with constant amplitude pulse trains. We found a significant correlation between 50-Hz BW NoSπ thresholds and ITD JNDs in experiment 1 [Fig. 2(A)], where we primarily tested at the most binaurally sensitive electrode pair. Therefore it is possible that this significant correlation is representing basic binaural sensitivity and underlying neural survival. However, we did not find a significant correlation between $\rho_{ref} = 1$ correlation change and ITD JNDs [Fig. 2(B)]. This may be because we had fewer CI listeners participate in this experiment, and the listeners were tested at multiple electrode pairs that had a range of binaural sensitivities. Therefore the large within-listener variability in experiment 2 may have produced the poor correlation between ITD and $\rho_{ref} = 1$ JNDs. The large within-listener variability may have resulted from the many psychometric functions that did not achieve 100% correct for large values of α. The relatively large $\rho_{ref} = 1$ JNDs would have been more susceptible to measurement noise. Note that previous NH studies have shown weak correlations between ITD JNDs and NoSπ or correlation change JNDs for non-vocoded stimuli (Koehnke et al., 1986; Koehnke et al., 1995). Therefore it seems likely that neural degeneration affected CI performance in this study.

In summary, further research is needed to identify the factors that are responsible for the performance difference between the CI and NH listeners. Possibilities include (a) a signal-based problem of imprecise envelope modulation encoding in CI listeners, (b) an overestimation of CI performance from the acoustic pulse trains presented to the NH listeners, (c) individual factors like neural degeneration or electrode array placement in the CI listeners, or (d) a combination of the above-mentioned factors.

C. Implications for speech processing strategies and speech understanding in noise

Clinical speech processors are not synchronized, meaning that they likely introduce distortions to ITDs and ILDs. In other words, clinical speech processors introduce their own interaural decorrelation. For example, automatic gain controls that lack across-ear synchronization would affect envelope modulation encoding and introduce ILD fluctuations. Decorrelation might also be introduced by compression functions. The implementation of the compression in Eq. (2) followed that used in previous CI research (e.g., Long et al., 2006),[6] which is an approximation of that used in speech processing strategies. The compression is often the same across all electrodes, and across both ears. However, it seems unlikely that loudness growth can be accommodated by the same compression when considering how sensitive humans are to binaural differences. Note that the largest amounts of interaural decorrelation as well as the largest instantaneous ITDs and ILDs occur in the temporal valleys of the signals (Buss et al., 2003; Goupell, 2012). Thus the processor DR becomes a relevant factor in representing interaural envelope correlation because it determines the depths of valleys. The temporal valleys are not represented with full control at precisely when the interaural differences in a signal are most prevalent and may be most important for detecting interaural decorrelation.

The experiments in this study used stimulation from single pairs of electrodes. In more realistic stimulation patterns (i.e., multi-electrode stimulation as would occur in a speech processor), current spread from adjacent electrodes would also add an uncontrolled change in interaural correlation. Tone-in-noise detection studies with stimulation on multiple electrode pairs in each ear show much smaller or no binaural unmasking. When multiple electrodes each presented NoSo or NoSπ, Van Deun et al. (2011) measured an average BMLD of 3 dB. When only one pair of electrodes presented NoSo or NoSπ and the other pairs presented No, Lu et al. (2011) measured that BMLDs were significantly reduced from 8.9 to 2.1 dB. Lu et al. also showed that the amount of current that spreads from the electrodes presenting No to the electrode region presenting NoSo or NoSπ was negatively correlated with magnitude of the BMLDs. The relatively poor BMLDs for multi-electrode stimuli in CI listeners is in contrast to NH listeners performing similar tasks, who maintain large BMLDs for large BW stimuli (van de Par and Kohlrausch, 1999).

In an attempt to transition from studying binaural unmasking of tones to using speech signals, van Hoesel et al. (2008) measured binaural unmasking of tones and speech with NoS($\tau = 0.7$ ms) for four bilateral CI listeners using synchronized research processors. They tested two speech processing strategies, one that discarded fine-structure timing cues and one that explicitly encoded fine-structure timing cues, which was called peak-derived timing. They found no binaural unmasking of tones or speech for any of the strategies. Therefore the lack of binaural unmasking in bilateral CI listeners cannot be attributed to the lack of encoding envelope ITDs because they used synchronized processors, and it cannot be attributed to the lack of fine-structure ITDs because they were presented in the peak-derived timing speech processing strategy. However, it may have been that the peak-derived timing strategy did not present the ITDs in a salient manner. Had the number of channels been reduced to limit channel interactions and had the apical channels been given coherent fine-structure timing, the low-frequency ITDs may have been salient and produced binaural unmasking (Churchill et al., 2014).

Loizou et al. (2009) also tested if binaural unmasking of speech stimuli could occur in bilateral CI listeners who were presented stimuli over synchronized research processors. They measured speech understanding for talkers that were co-located or spatially separated virtually with NH head-related transfer functions (HRTFs) (i.e., the listeners were provided both ITDs and ILDs in the envelope). They measured a 2–5 dB benefit for the virtually separated talkers compared to the co-located talkers. This benefit is much smaller than that found for NH listeners, which can be as large as 12 dB (e.g., Hawley et al., 2004). Loizou et al. concluded that the benefit for the bilateral CI listeners was mostly due to the monaural better-ear effect (i.e., the availability of a more favorable SNR at one of the ear). The lack of binaural unmasking in that study, even when using time-synchronized processors and pinna HRTFs rather than the

BTE microphone HRTFs, might be explained by interaural decorrelation. The speech signals from spatially separated talkers may not have been perceived at different locations because of the broadened or diffuse sound images from the uncontrolled interaural decorrelation produced by modulations and multiple electrode stimulation.

Given these issues, uncontrolled interaural correlation may be a major limitation for deriving larger binaural benefits in sound localization and spatial release from masking tasks in bilateral CI users. Solutions would include mapping the processors to respect the binaural cues at all levels between T and C (Goupell *et al.*, 2013), reducing channel interactions from spread of current with more focused stimulation, and utilizing stimulation strategies that explicitly attempt to not have stimulation from adjacent electrodes (i.e., a modification of the peak-picking speech encoding strategies). A sizeable reduction in the envelope correlation change JNDs would be necessary to have many of the CI listeners perform in the range of NH listeners, particularly for uncorrelated references. For spatial listening in rooms, nearby talkers (the ones that a listener wants to hear) have relatively correlated signals between the ears with only long-term ITDs or ILDs in the signal. However, the reflections and reverberation from the room and other distant background noise would arrive at the ears relatively interaurally decorrelated. Therefore it is critically important to have low $\rho_{ref} = 0$ JNDs to be able to distinguish more correlated and interesting foreground information from decorrelated background noise. The $\rho_{ref} = 0$ JNDs in our CI listeners with direct stimulation were mostly not determinable (Fig. 7). Therefore it would be difficult for these listeners to practically use the interaural envelope correlation to distinguish foreground talkers from background noise.

ACKNOWLEDGMENTS

[1]Note that we did not measure bilateral loudness growth functions, so it is an assumption that the compression function produced similar loudness growth functions across the ears in bilateral CI listeners. However, it is also an assumption that NH listeners experience similar loudness growth functions across the ears and if NH listeners have relatively more symmetric loudness growth functions compared to bilateral CI listeners.

[2]There was one non-monotonic psychometric function measured. For this psychometric function, only points on the monotonically increasing part of the function were used in the threshold calculation.

[3]Listener IBJ was unusual as she was not able to achieve above chance performance for either the NoSo or NoSπ detection tasks. For the ITD and NoSπ tasks, she reported that she did not perceive a fused binaural image but two sound sources, one at each ear. For the NoSo detection task, she reported that she could not detect changes in loudness or roughness consistently. Her DRs (defined as C minus T level) were about half of that for typical listeners (10–20 CUs), and the amplitudes needed to elicit T and C levels were substantially lower than our other listeners; this may have affected her ability to perform the NoSo detection task. Hence listener IBJ was omitted from all subsequent plots and analyses.

[4]It could also be the case that listeners were focusing so much on the binaural cues, that the monaural cues were ignored.

[5]Listener IBL was omitted because she showed no binaural sensitivity. Listener IBN began testing but could not perform above chance in detecting changes in interaural envelope correlation and so was also omitted.

[6]Note that this is an older approximation for Cochlear-type speech processing (the instantaneous processor DR is now 40 dB, not 30 dB) and that other manufacturers have different CU-μA functions and instantaneous processor DRs.

Bernstein, L. R., and Trahiotis, C. (**1996**). "On the use of the normalized correlation as an index of interaural envelope correlation," J. Acoust. Soc. Am. **100**, 1754–1763.

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.

Bronkhorst, A. W., and Plomp, R. (**1988**). "The effect of head-induced interaural time and level differences on speech intelligibility in noise," J. Acoust. Soc. Am. **83**, 1508–1516.

Brughera, A., Dunai, L., and Hartmann, W. M. (**2013**). "Human interaural time difference thresholds for sine tones: The high-frequency limit," J. Acoust. Soc. Am. **133**, 2839–2855.

Buss, E., Hall, J. W. 3rd, and Grose, J. H. (**2003**). "The masking level difference for signals placed in masker envelope minima and maxima," J. Acoust. Soc. Am. **114**, 1557–1564.

Churchill, T., Kan, A., Goupell, M. J., and Litovsky, R. Y. (**2014**). "Spatial hearing benefits demonstrated with presentation of acoustic temporal fine structure cues in bilateral cochlear implant listeners," J. Acoust. Soc. Am. **136**, 1246–1256.

Culling, J. F., Colburn, H. S., and Spurchise, M. (**2001**). "Interaural correlation sensitivity," J. Acoust. Soc. Am. **110**, 1020–1029.

Culling, J. F., Hawley, M. L., and Litovsky, R. Y. (**2004**). "The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources," J. Acoust. Soc. Am. **116**, 1057–1065.

Culling, J. F., Jelfs, S., Talbert, A., Grange, J. A., and Backhouse, S. S. (**2012**). "The benefit of bilateral versus unilateral cochlear implantation to speech intelligibility in noise," Ear Hear. **33**, 673–682.

Gabriel, K. J., and Colburn, H. S. (**1981**). "Interaural correlation discrimination: I. Bandwidth and level dependence," J. Acoust. Soc. Am. **69**, 1394–1401.

Goupell, M. J. (**2010**). "Interaural fluctuations and the detection of interaural incoherence. IV. The effect of compression on stimulus statistics," J. Acoust. Soc. Am. **128**, 3691–3702.

Goupell, M. J. (**2012**). "The role of envelope statistics in detecting changes in interaural correlation," J. Acoust. Soc. Am. **132**, 1561–1572.

Goupell, M. J., Kan, A., and Litovsky, R. Y. (**2013**). "Typical mapping procedures can produce non-centered auditory images in bilateral cochlear-implant users," J. Acoust. Soc. Am. **133**, EL101–EL107.

Goupell, M. J., and Litovsky, R. Y. (**2014**). "The effect of interaural fluctuation rate on correlation change discrimination," J. Assoc. Res. Otolaryngol. **15**, 115–129.

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (**2004**). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," J. Acoust. Soc. Am. **115**, 833–843.

Kan, A., Stoelb, C., Litovsky, R. Y., and Goupell, M. J. (**2013**). "Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant users," J. Acoust. Soc. Am. **134**, 2923–2936.

Koehnke, J., Colburn, H. S., and Durlach, N. I. (**1986**). "Performance in several binaural-interaction experiments," J. Acoust. Soc. Am. **79**, 1558–1562.

Koehnke, J., Culotta, C. P., Hawley, M. L., and Colburn, H. S. (**1995**). "Effects of reference interaural time and intensity differences on binaural performance in listeners with normal and impaired hearing," Ear Hear. **16**, 331–353.

Laback, B., and Majdak, P. (**2008**). "Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates," Proc. Natl. Acad. Sci. U.S.A. **105**, 814–817.

Lavandier, M., and Culling, J. F. (**2010**). "Prediction of binaural speech intelligibility against noise in rooms," J. Acoust. Soc. Am. **127**, 387–399.

Litovsky, R., Parkinson, A., Arcaroli, J., and Sammeth, C. (**2006**). "Simultaneous bilateral cochlear implantation in adults: A multicenter clinical study," Ear Hear. **27**, 714–731.

Litovsky, R. Y., Goupell, M. J., Godar, S., Grieco-Calub, T., Jones, G. L., Garadat, S. N., Agrawal, S., Kan, A., Todd, A., Hess, C., and Misurelli, S. (**2012**). "Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory," J. Am. Acad. Audiol. **23**, 476–494.

Litovsky, R. Y., Jones, G. L., Agrawal, S., and van Hoesel, R. (**2010**). "Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans," J. Acoust. Soc. Am. **127**, 400–414.

Litovsky, R. Y., Parkinson, A., and Arcaroli, J. (**2009**). "Spatial hearing and speech intelligibility in bilateral cochlear implant users," Ear Hear. **30**, 419–431.

Loizou, P. C. (**2006**). "Speech processing in vocoder-centric cochlear implants," in *Cochlear and Brainstem Implants*, edited by A. Moller (Karger, Basel, Switzerland), pp. 109–143.

Loizou, P. C., Hu, Y., Litovsky, R., Yu, G., Peters, R., Lake, J., and Roland, P. (**2009**). "Speech recognition by bilateral cochlear implant users in a cocktail-party setting," J. Acoust. Soc. Am. **125**, 372–383.

Long, C. J., Carlyon, R. P., and Litovsky, R. (**2007**). "Binaural unmasking with 'transposed' stimuli in bilateral cochlear implant users," in *Conference on Implantable Auditory Prostheses* (Asilomar, CA).

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.

Long, C. J., Eddington, D. K., Colburn, H. S., and Rabinowitz, W. M. (**2003**). "Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user," J. Acoust. Soc. Am. **114**, 1565–1574.

Lu, T., Litovsky, R., and Zeng, F. G. (**2010**). "Binaural masking level differences in actual and simulated bilateral cochlear implant listeners," J. Acoust. Soc. Am. **127**, 1479–1490.

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.

McKay, C. M., McDermott, H. J., and Clark, G. M. (**1994**). "Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees," J. Acoust. Soc. Am. **96**, 2664–2673.

Noel, V. A., and Eddington, D. K. (**2013**). "Sensitivity of bilateral cochlear implant users to fine-structure and envelope interaural time differences," J. Acoust. Soc. Am. **133**, 2314–2328.

Poon, B. B., Eddington, D. K., Noel, V., and Colburn, H. S. (**2009**). "Sensitivity to interaural time difference with bilateral cochlear implants: Development over time and effect of interaural electrode spacing," J. Acoust. Soc. Am. **126**, 806–815.

Rice, S. O. (**1954**). "Mathematical analysis of random noise," in *Selected Papers on Noise and Stochastic Processes*, edited by N. Wax (Dover, New York), pp. 133–294.

van de Par, S., and Kohlrausch, A. (**1999**). "Dependence of binaural masking level differences on center frequency, masker bandwidth, and interaural parameters," J. Acoust. Soc. Am. **106**, 1940–1947.

Van Deun, L., van Wieringen, A., Francart, T., Buchner, A., Lenarz, T., and Wouters, J. (**2011**). "Binaural unmasking of multi-channel stimuli in bilateral cochlear implant users," J. Assoc. Res. Otolaryngol. **12**, 659–670.

Van Deun, L., van Wieringen, A., Francart, T., Scherf, F., Dhooge, I. J., Deggouj, N., Desloovere, C., Van de Heyning, P. H., Offeciers, F. E., De Raeve, L., and Wouters, J. (**2009**). "Bilateral cochlear implants in children: Binaural unmasking," Audiol. Neurootol. **14**, 240–247.

van Hoesel, R. J. M. (**2012**). "Contrasting benefits from contralateral implants and hearing aids in cochlear implant users," Hear Res. **288**, 100–113.

van Hoesel, R. J. M., Bohm, M., Pesch, J., Vandali, A., Battmer, R. D., and Lenarz, T. (**2008**). "Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies," J. Acoust. Soc. Am. **123**, 2249–2263.

van Hoesel, R. J. M., Jones, G. L., and Litovsky, R. Y. (**2009**). "Interaural time-delay sensitivity in bilateral cochlear implant users: Effects of pulse rate, modulation rate, and place of stimulation," J. Assoc. Res. Otolaryngol. **10**, 557–567.

Whitmer, W. M., Seeber, B. U., and Akeroyd, M. A. (**2012**). "Apparent auditory source width insensitivity in older hearing-impaired individuals," J. Acoust. Soc. Am. **132**, 369–379.

Wichmann, F. A., and Hill, N. J. (**2001**). "The psychometric function: I. Fitting, sampling, and goodness of fit," Percept. Psychophys. **63**, 1293–1313.

Wightman, F. L., and Kistler, D. J. (**1992**). "The dominant role of low-frequency interaural time differences in sound localization," J. Acoust. Soc. Am. **91**, 1648–1661.

Yost, W. A., and Dye, R. H., Jr. (**1988**). "Discrimination of interaural differences of level as a function of frequency," J. Acoust. Soc. Am. **83**, 1846–1851.