

# Spatial hearing benefits demonstrated with presentation of acoustic temporal fine structure cues in bilateral cochlear implant listeners<sup>a)</sup>

Tyler H. Churchill and Alan Kan

Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, Wisconsin 53705

Matthew J. Goupell

Department of Hearing and Speech Sciences, University of Maryland-College Park, College Park, Maryland 20742

Ruth Y. Litovsky<sup>b)</sup>

Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, Wisconsin 53705

(Received 8 December 2013; revised 27 May 2014; accepted 14 July 2014)

Most contemporary cochlear implant (CI) processing strategies discard acoustic temporal fine structure (TFS) information, and this may contribute to the observed deficits in bilateral CI listeners' ability to localize sounds when compared to normal hearing listeners. Additionally, for best speech envelope representation, most contemporary speech processing strategies use high-rate carriers ( $\geq 900$  Hz) that exceed the limit for interaural pulse timing to provide useful binaural information. Many bilateral CI listeners are sensitive to interaural time differences (ITDs) in low-rate ( $< 300$  Hz) constant-amplitude pulse trains. This study explored the trade-off between superior speech temporal envelope representation with high-rate carriers and binaural pulse timing sensitivity with low-rate carriers. The effects of carrier pulse rate and pulse timing on ITD discrimination, ITD lateralization, and speech recognition in quiet were examined in eight bilateral CI listeners. Stimuli consisted of speech tokens processed at different electrical stimulation rates, and pulse timings that either preserved or did not preserve acoustic TFS cues. Results showed that CI listeners were able to use low-rate pulse timing cues derived from acoustic TFS when presented redundantly on multiple electrodes for ITD discrimination and lateralization of speech stimuli.

© 2014 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4892764>]

PACS number(s): 43.66.Qp, 43.66.Ts, 43.71.Ky [JFC]

Pages: 1246–1256

## I. INTRODUCTION

Cochlear implants (CIs) have provided hearing to hundreds of thousands of people worldwide who have severe-to-profound hearing loss. This technology has progressed from the single-electrode implant, which merely provided lip reading cues and sensation of sound, to present-day multi-channel devices, which can give some users excellent open-set speech recognition in quiet without lip reading. A timely question is whether there are feasible approaches that will enable the closing of the performance gap between CI users and normal hearing (NH) listeners. Electric hearing is not equivalent to normal hearing because the activation of the auditory nerve by CI electrode currents is not faithful to the pattern produced by normal physiological mechanisms (e.g., Moore, 2003). Notably, electric stimulation lacks the spatial and temporal resolution of NH mechanisms (Rubinstein and Miller, 1999), factors that are known to contribute to the degraded ability of CI users to understand speech in noise and localize sounds. Therefore, improving the spatial and temporal resolution of CI electric stimulation should improve users' hearing abilities.

A commonly used signal decomposition technique is to separate narrowband sounds' temporal envelopes from temporal fine structure (TFS) (e.g., Smith *et al.*, 2002). It has been found that speech is remarkably robust under signal degradation, and that speech understanding in quiet requires only a few spectrally contiguous channels of envelope information (Shannon *et al.*, 1995). Cochlear implants, with their relatively small number of available independent perceptual channels (about seven, Friesen *et al.*, 2001), have exploited that finding, representing sound envelopes by amplitude-modulating electrical pulse train carriers and discarding TFS information. Cochlear implant users have thus enjoyed much success for the case of speech in quiet, but less so for speech in noise and sound localization (van Hoesel and Tyler, 2003).

Much of NH listeners' ability to localize sounds relies heavily on the interaural time difference (ITD) information carried by acoustic TFS at frequencies below 1.5 kHz (Wightman and Kistler, 1992; Macpherson and Middlebrooks, 2002; Brughera *et al.*, 2013). Additionally, the ability of NH listeners to segregate target talkers and competing maskers by location is known to be important for understanding speech in noise (Bronkhorst and Plomp, 1988; Hawley *et al.*, 2004). This phenomenon, known as spatial release from masking (SRM), also largely depends on receiving ITDs carried by low-frequency TFS (Kidd *et al.*, 2010). With electric hearing, in the absence of TFS information, interaural level differences

<sup>a)</sup>Portions of this work were previously presented in poster form at the Conference on Implantable Auditory Prostheses, Pacific Grove, CA, July 2013.

<sup>b)</sup>Author to whom correspondence should be addressed. Electronic mail: litovsky@waisman.wisc.edu

(ILDs) and envelope ITDs are the only binaural cues available to bilateral CI listeners for spatially identifying sounds. These cues alone may be inadequate to achieve spatial release from masking (van Hoesel and Tyler, 2003; Ihlefeld and Litovsky, 2012). Therefore CI listeners' difficulty in listening conditions other than speech in quiet is likely due in part to the discarding of acoustic TFS by most of today's CI processing strategies, and it may be beneficial to reproduce acoustic TFS information in CIs' electrical signals.

In the normal auditory system, TFS is thought to be encoded in the timing of nerve firings that "phase lock" to mechanical oscillations of the basilar membrane produced by a sound. At electrical stimulation rates below several thousand pulses per second, a CI's current pulses induce "super" phase locking, whereby the auditory nerves fire synchronously with every current pulse (Kiang and Moxon, 1972). Therefore, acoustic TFS should be able to be accurately encoded by timing CI stimulating pulses to a given phase of the acoustic signal. Currently in CIs, constant- and high-rate ( $\geq 900$  Hz) pulse trains are typically used to represent speech envelopes as faithfully as possible (Loizou *et al.*, 2000; Galvin and Fu, 2005). While these rates are within the range of phase-locking of the auditory nerve to electric stimulation, they are beyond the range of pulse timing sensitivity for ITD (Majdak *et al.*, 2006; van Hoesel *et al.*, 2009) and rate discrimination (Carlyon *et al.*, 2010; but see Goldsworth and Shannon, 2014). Introduction of low-rate ( $< 300$  Hz) TFS information in the current pulse timing could enable the preservation of low-frequency TFS ITD cues, which may be critical for improved sound localization and speech understanding in noise by producing SRM.

Besides the loss of TFS information, there may be additional negative consequences to using constant-rate pulse trains for electrical stimulation in CIs. Supernatural phase-locking of the auditory nerve to constant-rate pulse trains impedes the natural stochasticity in the timing of nerve firings. Temporal jitter may be important to some NH mechanisms responsible for dynamic range, and there have been attempts to restore it (Rubinstein *et al.*, 1999; Chatterjee and Robert, 2001; Litvak *et al.*, 2001). Furthermore, the use of constant-rate pulse trains in CIs may lead to adaptation, wherein the nervous system disregards a repetitive signal, reducing the signal's information transfer capacity (Smith, 1979; Laback and Majdak, 2008).

There have been recent advances in CI processing that include attempts to address these deficits by timing the stimulating current pulses on a given electrode to represent acoustic TFS information in that particular channel (Zierhofer, 2003; van Hoesel, 2004; Nie *et al.*, 2005; Sit *et al.*, 2007). Some listeners occasionally report preference for strategies that deliberately time the individual pulses to follow fine-structure over conventional constant-rate strategies. However, following their take-home familiarization periods, results have not shown evidence of any significant benefits relative to other clinical strategies (Arnoldner *et al.*, 2007; Riss *et al.*, 2009; Schatzer *et al.*, 2010; Vermeire *et al.*, 2010). Additionally, although sensitivity to ITDs has been measured in bilaterally-implanted CI listeners (van Hoesel and Clark, 1997; van Hoesel *et al.*, 2009; Litovsky

*et al.*, 2010; Litovsky *et al.*, 2012), several studies that have investigated binaural benefits due to pulse timing derived from speech signals' acoustic TFS show no clear benefits of TFS-timed pulses (van Hoesel and Tyler, 2003; van Hoesel *et al.*, 2008). However, previous investigations into pulse timing sensitivity may have been confounded by the effects of current spread and the inclusion of ILDs. The present study used direct stimulation techniques and explicitly excluded ILD cues, leaving pulse timing and envelope ITDs as the only available cues for localization. With the goal of identifying a set of processing parameters that allow for both speech understanding and ITD sensitivity, we systematically examined the effects of channel pulse rate and pulse timing on ITD discrimination, ITD lateralization, and speech recognition of multi-channel speech stimuli. Results from these tests, conducted in quiet, may aid in the pursuit of better localization and understanding of speech in noise with bilateral CIs.

Given that ongoing ITD cues are most salient at low rates and ITD sensitivity diminishes at pulse rates above 300 Hz, and that high-rate pulse trains are superior at faithfully representing speech envelopes, a trade-off exists between speech understanding and ITD sensitivity. The present study has examined this trade-off by testing speech understanding and ITD sensitivity using direct stimulation with eight-channel speech stimuli at three electrical stimulation rate combinations: (1) low rates (100–173 Hz) on all channels, (2) low rates on four apical channels and high rates (894–1547 Hz) on four basal channels, and (3) high rates on all channels. A novel TFS-retaining strategy (Churchill, 2014) was tested against a conventional strategy, continuous interleaved sampling (CIS) (Wilson *et al.*, 1991). By testing with both the TFS and CIS strategies, we were able to compare listeners' sensitivity to envelope and TFS ITDs across the three rate combinations. In order to test exclusively the effect of pulse timing, for a given rate combination and stimulus token, the TFS and CIS strategies were designed such that each produced a stimulus with the same average pulse rate.

## II. METHODS

### A. Experimental protocols

Eight bilateral CI users (see Table I) each performed the following three tasks: closed-set speech recognition using four- or five-word sentences, lateralization of single-word stimuli, and left/right discrimination of single-word stimuli. All listeners normally wore Cochlear Ltd. (Sydney, Australia) devices using the Advanced Combination Encoder (ACE) speech strategy at a pulse rate  $\geq 900$  Hz. However, for these experiments, listeners used special bilaterally-synchronized research processors (Cochlear L34 speech processors) that were attached to a personal computer and controlled by custom software run in MATLAB (the Mathworks, Natick, MA). Each of these listeners had demonstrated ITD sensitivity in previous experiments at the University of Wisconsin-Madison.

Listener responses were collected via a graphical user interface on a touchscreen, and correct-answer visual feedback was provided following each trial. Two processing

TABLE I. Listener data and electrodes used for testing.

| Subject | Age | Hearing aid Use (yr) | L/R CI Use (yr) | Left electrodes      | Right electrodes      |
|---------|-----|----------------------|-----------------|----------------------|-----------------------|
| IAJ     | 68  | 46                   | 17/9            | 1,3,6,9,13,15,18,20  | 3,5,8,11,14,18,20,22  |
| IBF     | 61  | 14                   | 5/7             | 2,4,6,8,12,15,18,21  | 3,6,8,10,13,16,19,22  |
| IBK     | 73  | 8                    | 10/4            | 2,4,6,9,12,15,18,20  | 4,7,8,11,14,17,20,22  |
| IBM     | 59  | 16                   | 3/7             | 2,4,7,10,13,16,19,22 | 2,4,7,10,13,16,19,22  |
| IBN     | 66  | 50                   | 3/13            | 2,4,6,8,10,13,18,20  | 6,8,10,12,14,17,20,22 |
| IBR     | 59  | 22                   | 3/9             | 2,4,7,11,15,17,19,21 | 2,4,7,11,13,15,17,19  |
| ICD     | 56  | 40                   | 4/10            | 4,6,9,12,15,17,20,22 | 2,4,7,10,13,15,18,20  |
| ICM     | 60  | 29                   | 3/1             | 2,4,6,9,12,15,18,20  | 3,5,7,11,14,17,20,22  |

strategies and three stimulation rate combinations were tested in every task ( $2 \times 3 = 6$  total conditions). For the speech recognition task, a given sentence was constructed using one of the six permutations of rate combination and strategy. The presentation order of the six permutation conditions was randomized among trials for all tasks. A test session for a given task lasted approximately 20 min, and the order of tasks conducted was interleaved so as to support listener alertness and engagement. Informed consent was obtained, and listeners were monetarily compensated for their participation. All procedures were approved by the University of Wisconsin's Human Subjects Institutional Review Board.

For the speech recognition task, listeners were asked to identify each of the four or five words in a low-context sentence, e.g., "Mike bought five cards," "Jill lost four red hats," etc. Sentences were those spoken by Male #1 from the Kidd *et al.* (2008) corpus. Target stimuli were presented from 0° azimuth, i.e., with an ITD of zero. All words were weighted identically in scoring percent correct. For the lateralization task, listeners identified the perceived azimuthal location of the presented stimulus. Lateralization stimuli consisted of the names from the above sentence corpus, and the applied ITDs were calculated from head-related transfer functions (HRTFs) of the KEMAR manikin (Algazi *et al.*, 2001) for source azimuths from  $-70^\circ$  to  $+70^\circ$  in  $10^\circ$  increments for most listeners, and from  $-50^\circ$  to  $+50^\circ$  in  $10^\circ$  increments and  $\pm 90^\circ$  for two earlier listeners (IAJ and IBF). For the discrimination task, listeners performed a two-interval, two-alternative forced-choice task in which they indicated whether the stimulus' perceived location moved from right-to-left or from left-to-right between two presentations of the same speech token. The second interval contained an ITD opposite of that contained in the first; listeners discriminated positive and negative ITDs of 50, 100, 200, 400, 800, and 1600  $\mu$ s. Discrimination stimuli also consisted of names from the sentence corpus used in the speech recognition task.

Prior to testing, new research maps were created for each listener. This was done to find threshold and comfortable current levels for all active electrodes in left and right implants at rates representative of low- and high-rate test stimuli: 150 and 1500 Hz, respectively. This mapping is in many ways similar to the procedure performed by the listeners' audiologists for their clinical devices. Stimulation consisted of 300-ms trains of monopolar, biphasic current pulses with phase durations of 25  $\mu$ s and an interphase interval of

8  $\mu$ s. Next, current amplitude levels were adjusted to produce equal loudness sensations for all electrodes at comfortable-level stimulation at both high and low rates. This was performed by sequentially activating five electrodes and asking the listener to indicate which intervals needed to be adjusted in order to make them all have the same perceived loudness. This procedure was repeated until all electrodes were at equal loudness. Finally, levels on each electrode pair were adjusted to produce ILD-centered auditory images. Electrode pairs were activated simultaneously, and listeners provided feedback to the experimenter as to whether the auditory image was centered or to the left or right, following which the experimenter would adjust the levels to produce a more centered auditory image. Because place-pitch matching of left/right electrodes has been found to be important for ITD sensitivity (van Hoesel *et al.*, 2008; Poon *et al.*, 2009; Kan *et al.*, 2013), eight binaurally pitch-matched electrode pairs were selected for the presentation of the test stimuli. Pitch-matched electrode pairs were selected based on data collected from these listeners in ongoing experiments in the lab, wherein the listeners performed direct pitch comparisons for numerous pairs of electrodes (Litovsky *et al.*, 2012). The electrode pairs used for each listener are shown in Table I.

For speech recognition in quiet, 100 words (20 or 25 sentence trials) were presented for each strategy and rate combination, resulting in a total of 600 individual word-strategy-rate combination trials. In order to minimize ceiling or floor effects, resulting percent correct scores were arcsine-transformed (Studebaker, 1985) prior to analyses. For ITD discrimination, at least 40 repetitions were collected on at least four-point ITD psychometric functions for each strategy and rate combination. Logistic function psychometric curves (Wichman and Hill, 2001) were fitted through ITD discrimination percent correct points to calculate just-noticeable differences (JNDs, 71% correct) for each listener and also for pooled response data. Some listeners showed no ITD discrimination sensitivity with the CIS strategy for several rate combinations, and these JNDs are reported as "not measurable." For ITD lateralization, 10 repetitions were collected for each strategy, rate combination, and azimuth. Linear best-fit psychometric functions (response azimuths as functions of input azimuths) were calculated for each listener and also for pooled responses. The slopes of these input-output functions characterize the listeners' ability to use the available cues and are used here as the primary metrics of lateralization ability.

## B. Processing strategies

Figure 1 shows a block diagram representing the signal processing steps described below. Following the application of ITDs between the left and right channels, stimuli were resampled from 44.1 to 100 kHz, and henceforth each channel was processed separately. This independence of left/right processing could allow for implementation on non-linked processors, but the assurance of controlled, synchronous stimulation would be tenuous. As described below, envelopes and TFS were extracted using separate techniques.

A short-time Fourier transform-based method was used for envelope extraction. Preliminary testing indicated that listeners found speech more intelligible using this method than when using rectify-and-filter time-domain methods. Stimuli were first buffered into 512-point (5.12 ms) slices with a 256-point (2.56 ms) overlap between adjacent time slices. Next, a 512-point Blackman window was applied on each time slice, and a 512-point Fourier transform was performed on each time slice. In order to obtain eight envelope channels between 200 Hz and 16 kHz, nine logarithmically-spaced bin corners were chosen. The signal envelopes were calculated by summing the spectral magnitudes within the logarithmic corners assigned to each channel at each time slice. The resulting envelope was up-sampled by a factor of 256 in order to match the time resolution of pulse timing information. This resulted in a 100 kHz-sampled envelope for each of the eight channels and for each ear.

Pulse timing information was extracted from the signal at each ear in two bands, a low-rate band (100–173 Hz) and a high-rate band (894–1547 Hz). It should be noted that these bands are unrelated to the channels used in the envelope extraction and are based on a logarithmic spacing of eight channels between 100 Hz and 8 kHz. The use of only two bands for calculating the pulse timing information provided redundancy across stimulation channels. Preliminary testing

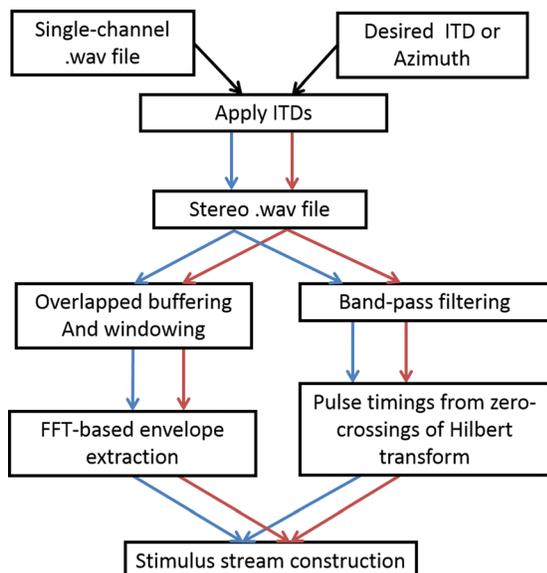


FIG. 1. (Color online) Block diagram for the TFS processing strategy. Rates used for the CIS strategy's isochronous stimulating pulse trains were the average rates of the pulse trains obtained for the corresponding speech token processed by the TFS strategy.

had indicated that redundant TFS cues were more salient than independent TFS cues calculated for each channel. Filtering was performed on the signal at each band with third-order Butterworth filters using forward-and-reverse filtering, a zero phase-shift technique which doubles the effective filter order. Next, a Hilbert transform was performed on the output of each of the two bands on each side to generate the analytic signal. Then, the positive-going zero-crossings of the phase of the analytic signal were extracted. This process yielded 100-kHz sampled vectors consisting of ones at the positive-going zero-crossings and zeroes elsewhere. The output of this processing consisted of one low-rate and one high-rate pulse timing vector for each ear, where the pulses were aligned to the peaks of the original signal at each band. We chose to derive the pulse timing information from the zero-crossings of the phase of the analytic signal rather than directly from the zero-crossings of the real signal because results from Monte Carlo simulations found that the zero-crossings of the real signal were roughly 4% more susceptible to timing corruption in the presence of additive white noise. The constant-rate pulse timing vectors for the corresponding CIS-processed stimuli were created by calculating the average pulse rate of each TFS-based timing vector, and creating constant-rate pulse timing vectors based on these average rates.

Pulse timing vectors (high- and/or low-rate, TFS-based or constant-rate CIS) were then modulated by the appropriate envelopes based on the strategy and rate combination to be presented on a given trial. For the low-rate stimuli, all eight stereo channels' envelopes modulated the corresponding left and right low-rate pulse timing vectors. For the high-rate stimuli, all eight channels' envelopes modulated the corresponding left and right high-rate pulse timing vectors. For the mixed-rate stimuli, the four apical (low frequency) channels' envelopes modulated the low-rate pulse timing vectors, and the four basal (high frequency) channels' envelopes modulated the high-rate pulse timing vectors. Modulated pulse timing vectors were then resampled into 70- $\mu$ s wide bins and power-law compressed (exponent = 1/3) into current levels between the listeners' threshold and comfort levels. Interleaving across electrodes was as follows for both TFS and CIS conditions (electrodes numbered 1 to 8 apical to basal): 1, 5, 2, 6, 3, 7, 4, 8. Figure 2 shows an example electrodiagram segment from a mixed-rate, TFS-strategy stimulus token for one ear. Stimuli were generated prior to listeners beginning the experiments. Table II shows stimulus pulse rate means and standard deviations for each task and rate combination.

## III. RESULTS

Figure 3(A) shows results from the speech recognition task in quiet. Listeners' average percent correct scores are shown for each strategy and rate combination. The right-most sets of bars also show averages across rates for each strategy, and across strategies for each rate combination. One-way, within-subject analyses of variance (ANOVAs) conducted for strategy, rate combination, and the interaction of strategy and rate combination found significant effects of

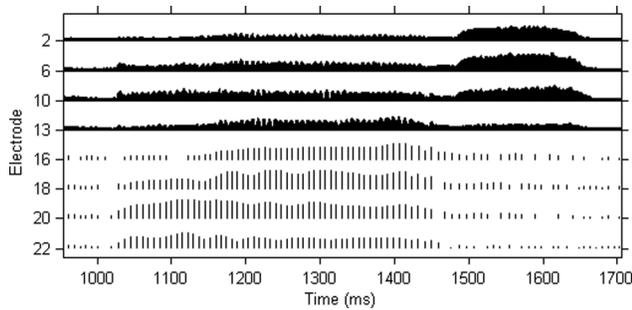


FIG. 2. Electrodiagram showing pulse timing and amplitudes for a segment of a mixed-rate, TFS strategy stimulus token. Envelopes of speech signals are presented bilaterally on eight pitch-matched electrode pairs. Pulse timing for the TFS strategy contains relevant ITD information, while all CIS strategy tokens are presented by constant rates with zero ITD.

rate combination [ $F_{(2,14)} = 12.8, p < 0.001$ ] and the interaction of strategy and rate combination [ $F_{(5,35)} = 10.2, p < 0.001$ ]. Combining the data for all listeners and rate combinations, planned *post hoc* Bonferroni-corrected two-tailed paired t-tests indicated that scores with high-rate stimuli were larger than with mixed ( $p = 0.012$ ) or low rates ( $p < 0.001$ ) and that scores with mixed stimulation rates were larger than scores with low rates ( $p = 0.002$ ). Planned, two-tailed, paired t-test of scores grouped by strategy indicated that scores with the TFS strategy were larger overall than those with the CIS strategy ( $p = 0.029$ ). In summary, high stimulation rates resulted in better speech understanding, while at low and mixed rates, TFS pulse timing appeared to contribute positively to speech understanding.

Figure 3(B) shows the averaged slopes of the lateralization psychometric function, and Fig. 4 shows two examples of individual listener average response data. Higher slopes indicate better sensitivity to ITDs. As can be seen from the lateralization curves plotted in Fig. 4, stimuli with low-rate, TFS-timed pulses produced the widest range of position responses and the largest psychometric function slopes, whereas stimuli with no pulse timing cues resulted in response curves with small slopes. Individual listeners' psychometric functions were analyzed to derive function slopes, correlation values, and significance levels, which are shown in Table III. Positive and significant slopes ( $p < 0.01$ ) were observed for all listeners with low- and mixed-rate TFS-processed stimuli, indicating that listeners were able to use low-rate pulse timing for lateralization. One-way, within-subjects ANOVAs were conducted on lateralization slopes for strategy, rate combination, and their interaction. These analyses revealed significant effects of strategy [ $F_{(1,7)} = 47.1, p < 0.001$ ], rate combination [ $F_{(2,14)} = 10.0, p = 0.002$ ], and

TABLE II. Stimulus pulse rate means (standard deviations) for given rate combinations and tests.

| Mean (s.d.) pulse rates |                | Rate combination |                   |           |
|-------------------------|----------------|------------------|-------------------|-----------|
|                         |                | Low              | Mixed             | High      |
| Test                    | Speech rec     | 129(1)           | 111(5)/1148(59)   | 1157(61)  |
|                         | Discrimination | 125(3)           | 108(10)/1093(163) | 1079(139) |
|                         | Lateralization | 131(3)           | 114(11)/1131(47)  | 1150(48)  |

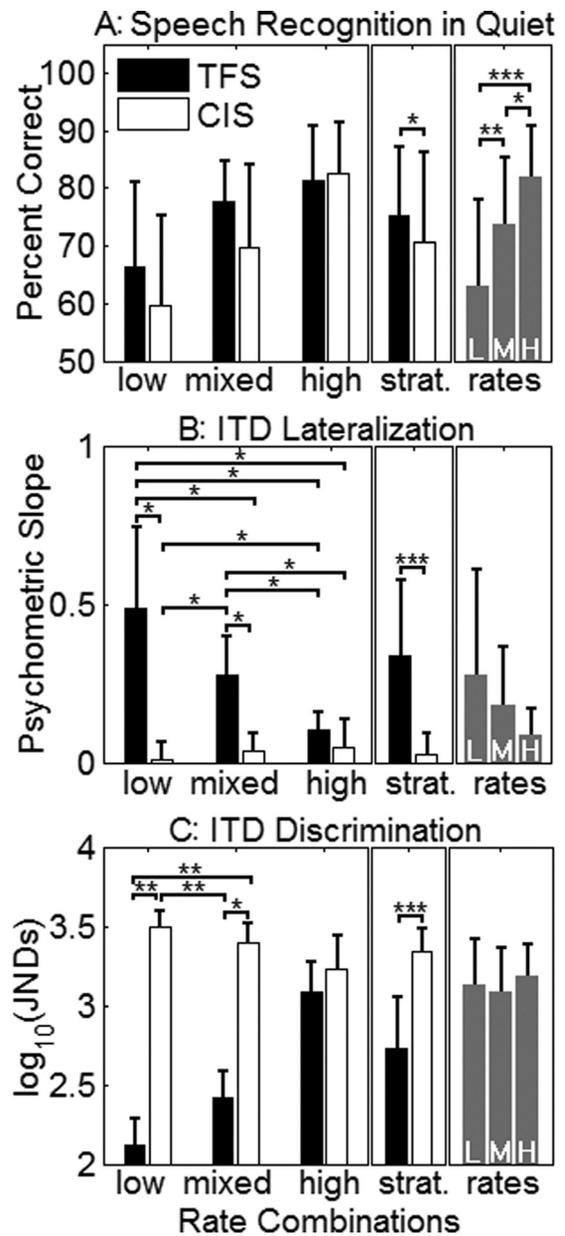


FIG. 3. Results from three tests, averaged across listeners for each rate combination and processing strategy. Within each panel, the three sets of bars on the left show metric averages for low, mixed, and high rates. The fourth pair of bars shows results averaged across all rate combinations for TFS and CIS strategies. The fifth, right-most set of bars shows metrics averaged across the two strategies for each of the three rates. Panel (A) shows percent correct scores for speech recognition in quiet. Panel (B) shows ITD lateralization psychometric function slopes. Panel (C) shows ITD discrimination JNDs. Error bars show the standard deviation of measurable listener scores. Asterisks indicate significant differences (\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ).

the interaction term [ $F_{(5,35)} = 27.2, p < 0.001$ ]. Planned *post hoc* Bonferroni-corrected two-tailed paired t-tests between all possible combinations of strategy and rates (15 total) revealed that low-rate and mixed-rate TFS-processed stimuli resulted in higher lateralization slopes than any other conditions ( $p < 0.05$ ) and that high-rate TFS-processed stimuli resulted in higher lateralization slopes than low-rate CIS-processed stimuli ( $p < 0.05$ ). A planned two-tailed paired t-test of lateralization slopes grouped by strategy indicated that slopes with the TFS strategy were larger overall than those with the

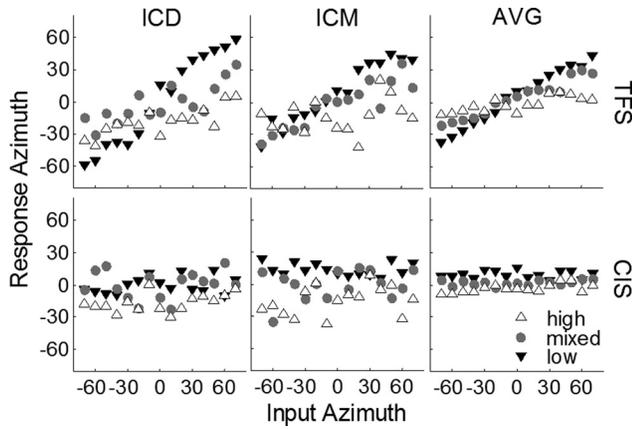


FIG. 4. Average lateralization response data for two listeners, ICD and ICM, and the averages across all listeners, shown for each strategy and rate combination. The top row shows responses with the TFS strategy and the bottom row shows responses with the CIS strategy.

CIS strategy ( $p < 0.001$ ). Planned Bonferroni-corrected paired t-tests of slopes grouped by rate combination revealed no differences. In summary, these findings indicate that listeners were able to use low- and mixed-rate TFS-timed pulses to lateralize speech stimuli.

TABLE III. Individual listeners' ITD lateralization psychometric function slopes, correlation values, and significance levels. Positive and significant slopes ( $p \leq 0.05$ ) are shown in bold.

| Subject | Rates | CIS          |              |              | TFS          |              |                  |
|---------|-------|--------------|--------------|--------------|--------------|--------------|------------------|
|         |       | Slope        | $R$          | $p$          | Slope        | $R$          | $p$              |
| IAJ     | low   | -0.003       | -0.018       | 0.945        | <b>0.328</b> | <b>0.880</b> | <b>&lt;0.001</b> |
|         | mixed | 0.024        | 0.109        | 0.676        | <b>0.200</b> | <b>0.673</b> | <b>0.003</b>     |
|         | high  | <b>0.141</b> | <b>0.527</b> | <b>0.030</b> | 0.015        | 0.062        | 0.815            |
| IBF     | low   | -0.030       | -0.308       | 0.305        | <b>0.272</b> | <b>0.866</b> | <b>&lt;0.001</b> |
|         | mixed | 0.035        | 0.325        | 0.278        | <b>0.204</b> | <b>0.841</b> | <b>&lt;0.001</b> |
|         | high  | 0.060        | 0.307        | 0.307        | 0.107        | 0.482        | 0.095            |
| IBK     | low   | -0.081       | -0.451       | 0.092        | <b>0.494</b> | <b>0.925</b> | <b>&lt;0.001</b> |
|         | mixed | 0.019        | 0.126        | 0.655        | <b>0.328</b> | <b>0.816</b> | <b>&lt;0.001</b> |
|         | high  | <b>0.192</b> | <b>0.597</b> | <b>0.019</b> | <b>0.167</b> | <b>0.514</b> | <b>0.050</b>     |
| IBM     | low   | 0.022        | 0.110        | 0.697        | <b>0.954</b> | <b>0.965</b> | <b>&lt;0.001</b> |
|         | mixed | 0.080        | 0.379        | 0.163        | <b>0.572</b> | <b>0.884</b> | <b>&lt;0.001</b> |
|         | high  | -0.034       | -0.329       | 0.231        | <b>0.115</b> | <b>0.514</b> | <b>0.050</b>     |
| IBN     | low   | -0.057       | -0.201       | 0.473        | <b>0.456</b> | <b>0.916</b> | <b>&lt;0.001</b> |
|         | mixed | -0.099       | -0.423       | 0.116        | <b>0.277</b> | <b>0.859</b> | <b>&lt;0.001</b> |
|         | high  | -0.031       | -0.228       | 0.414        | 0.104        | 0.449        | 0.093            |
| IBR     | low   | 0.088        | 0.238        | 0.393        | <b>0.378</b> | <b>0.664</b> | <b>0.007</b>     |
|         | mixed | 0.014        | 0.056        | 0.842        | <b>0.365</b> | <b>0.878</b> | <b>&lt;0.001</b> |
|         | high  | -0.085       | -0.400       | 0.140        | 0.124        | 0.389        | 0.151            |
| ICD     | low   | 0.053        | 0.300        | 0.277        | <b>0.922</b> | <b>0.981</b> | <b>&lt;0.001</b> |
|         | mixed | 0.040        | 0.136        | 0.629        | <b>0.316</b> | <b>0.801</b> | <b>&lt;0.001</b> |
|         | high  | 0.085        | 0.448        | 0.094        | <b>0.226</b> | <b>0.775</b> | <b>0.001</b>     |
| ICM     | low   | -0.027       | -0.209       | 0.455        | <b>0.602</b> | <b>0.967</b> | <b>&lt;0.001</b> |
|         | mixed | 0.084        | 0.272        | 0.326        | <b>0.434</b> | <b>0.904</b> | <b>&lt;0.001</b> |
|         | high  | 0.103        | 0.343        | 0.211        | 0.099        | 0.283        | 0.307            |
| AVG     | low   | -0.006       | -0.048       | 0.855        | <b>0.485</b> | <b>0.949</b> | <b>&lt;0.001</b> |
|         | mixed | <b>0.039</b> | <b>0.538</b> | <b>0.026</b> | <b>0.277</b> | <b>0.876</b> | <b>&lt;0.001</b> |
|         | high  | 0.045        | 0.260        | 0.314        | <b>0.104</b> | <b>0.659</b> | <b>0.004</b>     |

Figure 3(C) shows average JNDs for opposite-signed ITD pairs. At low and mixed rates, listeners had better (lower) JNDs with TFS-processed stimuli than with CIS-processed stimuli, but these JNDs converged at high rates. One-way, within-subjects ANOVAs were conducted on discrimination JNDs for strategy, rate combination, and their interaction. These analyses revealed significant effects of strategy [ $F_{(1,7)} = 66.3, p < 0.001$ ] and the interaction term [ $F_{(5,35)} = 17.6, p < 0.001$ ]. Planned *post hoc* Bonferroni-corrected two-tailed paired t-tests between all combinations of strategy and rates revealed that low-rate, TFS-processed stimuli produced lower discrimination JNDs than either low-rate or mixed-rate CIS-processed stimuli ( $p < 0.01$ ), and that mixed-rate, TFS-processed stimuli produced lower JNDs than either low-rate ( $p < 0.01$ ) or mixed-rate ( $p < 0.05$ ) CIS-processed stimuli. A planned two-tailed paired t-test of JNDs grouped by strategy indicated that JNDs with the TFS strategy were lower overall than those with the CIS strategy ( $p < 0.001$ ). T-tests of JNDs grouped by rate combination revealed no differences. In summary, listeners were able to use low- and mixed-rate TFS-timed pulses to discriminate ITDs in speech stimuli.

Overall, listeners performed better or equally with TFS-processed stimuli versus CIS-processed stimuli on all tasks. Compared to results obtained with the CIS strategy, the TFS strategy showed improved discrimination and lateralization scores without significant detriment to speech scores. Figure 5 shows relationships among the three measures whose averages are plotted in Fig. 3. Figure 5(A) demonstrates the relationship between lateralization and discrimination ability; larger slopes correlated with lower JNDs ( $R^2 = 0.454, p < 0.001$ ). Best overall spatial hearing ability, i.e., lower discrimination JNDs and larger lateralization slopes, is indicated by data points in the lower-right corner of the plot. Figures 5(B) and 5(C) illustrate the trade-off between ITD sensitivity at low rates and superior speech recognition at high rates. In Fig. 5(B), the ideal combination of scores (high lateralization slopes and speech recognition scores) is indicated by points approaching the upper-right corner. In Fig. 5(C), low JNDs and high speech recognition percent correct are indicated by points approaching the lower-right corner of the plot. These "ideal" performance regions are populated by only a few data points, e.g., IBM with low rates and IBF with mixed rates on Fig. 5(B). However, these points are exclusively from the low-rate and mixed-rate TFS-processed stimuli, and it is evident that overall, the TFS strategy more closely approaches these ideal performance regions than the CIS strategy. Even though there is large between-listener variability, a low- or mixed-rate TFS strategy may provide the optimal set of cues for spatial hearing and speech recognition in quiet for some listeners.

The listeners had no familiarization exposure to the stimuli prior to testing. In order to assess learning effects on the speech recognition task in quiet, the response data for each listener were organized chronologically and a cumulative correct function (CCF) was generated as a function of trial number  $t$  for each strategy/rate combination condition and for each strategy summed across rate combinations:

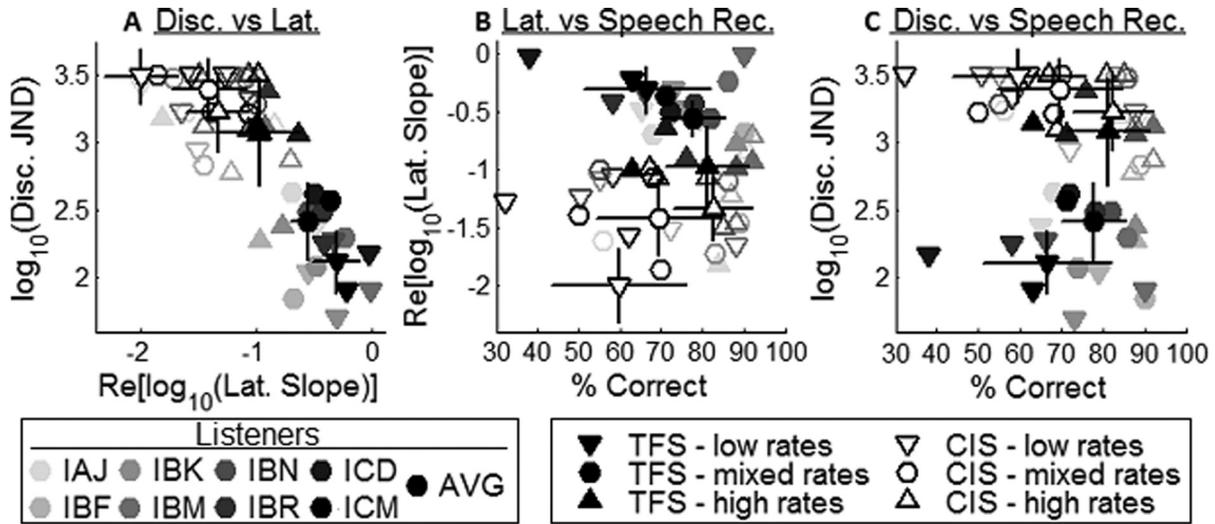


FIG. 5. Relationships among measures from all three tests. Grayscale markers indicate individual listeners' scores, while black markers with error bars indicate average scores from pooled listener response data. Error bars show standard deviation of measurable listener scores. Panel (A) shows ITD discrimination JNDs versus ITD lateralization psychometric function slopes, panel (B) shows lateralization slopes versus percent correct speech recognition in quiet, and panel (C) shows discrimination versus speech recognition.

$$CCF(t) = \sum_{t'=1}^t TC(t'),$$

where  $TC(t')$  (trial correct) is equal to one if the listener responded correctly on trial  $t'$  and zero otherwise. The following three models were used to fit the CCFs:

$$(A) CCF(t) = \beta_l \times t,$$

$$(B) CCF(t) = \beta_q \times t^2,$$

$$(C) CCF(t) = \beta_l \times t + \beta_q \times t^2,$$

where  $\beta_l$  and  $\beta_q$  are the linear and quadratic coefficients, respectively. Model A represents the CCF as a pure linear function in trial number  $t$ , and ignores any improvements in performance over time. Model B represents the CCF as a pure quadratic function in trial number and assumes no base level of understanding. Model C represents the CCF as a sum of linear and quadratic terms in trial number. It assumes a base of speech understanding would provide a linear increase in number correct with trial number and learning effects are represented by a quadratic term in trial number. Two-tailed t-tests of model coefficients found that both models A and C showed larger  $\beta_l$  values for the TFS strategy in the mixed-rate condition ( $P=0.0486$  and  $P=0.0036$ , respectively). These results suggest that there was no significant difference in rates of learning the two strategies, but that speech understanding was easier with the TFS strategy than the CIS strategy in the mixed-rate condition, a prediction reflected in the final percents correct for these conditions.

#### IV. DISCUSSION

The results reported here show that redundant, low-rate pulse timing on multiple channels can carry useful ITD

information for lateralization and discrimination of speech stimuli with bilateral CIs, even when mixed with high rates on some channels (see Figs. 3–5). The results also show that this useful pulse timing information can be calculated directly from the signals' TFS. Furthermore, the results characterize the trade-off between lateralization abilities and speech recognition in quiet as different pulse rates and processing strategies are used. From the depiction of this trade-off in panels (B) and (C) of Fig. 5, it is clear that the TFS-processed stimuli are better than CIS-processed stimuli for providing some listeners with both speech and localization cues. However, a single strategy and rate combination did not stand out as the universal best parameter set. While for one listener, the best strategy/rate combination was low-rate TFS, for another listener, the ideal combination may be mixed-rate TFS. Ideal parameter settings are not consistent across all CI users and listening conditions, but these results suggest that bilateral CI listeners may benefit from the inclusion of low-rate or multi-rate, TFS-timed strategies with their clinical maps. A minor issue confounding these results is that summation of loudness across electrodes may vary depending on stimulation strategy and rate. While most of the current knowledge on loudness summation is based on periodic stimulation, it is unclear how aperiodic stimulation affects loudness summation, both unilaterally and bilaterally. Efforts were made to reduce loudness differences across the different rate conditions prior to testing, and no subject reported appreciable loudness differences across the different rate conditions. Hence, we believe that performance was not largely affected by small loudness differences.

While it is apparent that the listeners tested here could lateralize stimuli due to their sensitivity to ITDs in the pulse timing, it is important to note that we deliberately ignored a dominant spatial cue for CI listeners, ILDs. Three listeners (IAJ, IBF, and IBK) also completed lateralization testing in which the stimuli contained the full set of lateral cues available in the KEMAR HRTFs, i.e., ITDs and ILDs. However,

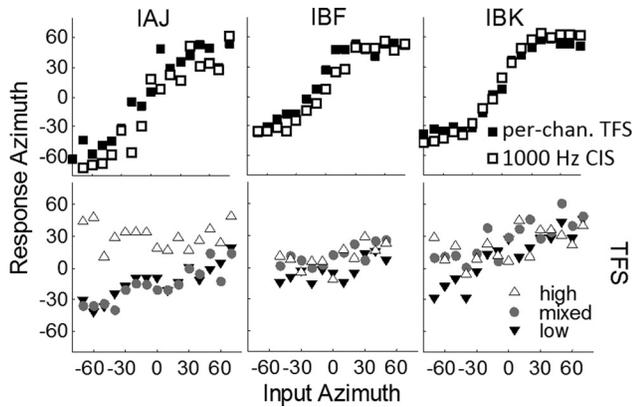


FIG. 6. Average lateralization responses from three listeners who completed lateralization of stimuli with ITD cues only (TFS strategy only, bottom row) and also with all HRTF cues (top row). The CIS strategy with full HRTF cues used 1000 Hz pulse trains on all electrodes, and the TFS strategy with full HRTF cues used pulse timing derived uniquely from the TFS of each channel.

the strategies used for this subset of testing were different from those tested in the other tasks; this CIS strategy used 1000-Hz constant-rate pulse trains on all channels, and this TFS strategy used pulse timings determined independently for each channel. Therefore, pulse timing cues may not have been as accessible with this “per-channel TFS” strategy. Average listener responses for these and the present study’s ITD-only lateralization tests are shown in Fig. 6. As can be seen by comparing the nearly indistinguishable sigmoidal response patterns for these two strategies (Fig. 6, top row), it appears that pulse timing ITDs did not contribute significantly to improving lateralization when both ITDs and ILDs were available. Comparing responses to stimuli with ITD and ILD cues (Fig. 6, top row) to responses to stimuli with only ITD cues (Fig. 6, bottom row), we see that the inclusion of ILD cues resulted in steeper slopes, but also produced the characteristic hemispherical sensitivity displayed by CI listeners in free-field studies. This pattern of hemispherical sensitivity might be broken by the inclusion of prominent and redundant low-rate ITD cues in the pulse timing.

Free-field and direct-connect studies have previously shown that for speech in noise, bilateral CI listeners benefit from having two implants in several ways (van Hoesel and Tyler, 2003; Schleich *et al.*, 2004; Litovsky *et al.*, 2006; Litovsky *et al.*, 2009; Loizou *et al.*, 2009). First, they may attend to their better ear, the ear with a more favorable SNR, taking advantage of monaural head shadow effects for 4–5 dB of improvement. Second, some listeners may also exhibit an additional 1–2 dB binaural benefit. Since speech TFS cues are normally unavailable, this binaural benefit is thought to be derived from ILDs and/or envelope ITDs. Whereas NH listeners show robust spatial release from masking with symmetrically separated maskers, other work has shown that bilateral CI users perform poorly on conditions in which maskers are symmetrically distributed and monaural head shadow cues are minimal or absent (Misurelli and Litovsky, 2012). Some studies have also compared listeners’ ILD and ITD sensitivities, finding that bilateral CI listeners are vastly more sensitive to natural ILD cues than

natural ITD cues (van Hoesel and Tyler, 2003; Litovsky *et al.*, 2010). Direct stimulation experiments by Long *et al.* (2006) and Lu *et al.* (2010) found that binaural unmasking of non-speech signals could be achieved with envelope decorrelation alone, but did not investigate the effects of TFS. Loizou *et al.* (2009) examined spatial release from masking in bilateral CI listeners by presenting listeners speech and informational masker stimuli through the auxiliary ports of bilaterally linked research processors, but found no binaural advantage for spatial release from masking with a conventional constant-rate strategy. The authors suggested that this lack of binaural advantage arises from poor ITD sensitivity, poor spectral resolution, and/or binaural mismatch.

The current study has attempted to avoid binaural mismatch, and the enhancement of ITD sensitivity with our TFS-processed stimuli may provide the necessary cues for spatial release from masking. van Hoesel *et al.* (2008) investigated the use of target ITDs for speech unmasking with several processing strategies, including one that represented per-channel TFS in pulse timing at several low-frequency apical electrodes. This strategy, peak-derived timing (PDT), preserves TFS cues by timing pulses to the positive peaks in the output of each channel’s filter (van Hoesel, 2004). The study found no binaural speech unmasking when applying a 700- $\mu$ s ITD to the target signal in the presence of a masker presented from the front (ITD = 0  $\mu$ s). That study also investigated free-field lateralization of click train stimuli, and found no significant differences among PDT, ACE, and CIS. However, the inclusion of ILD cues in free-field lateralization testing likely masked the impact of pulse timing ITDs. Thus, ILD cues may be more usable for binaural unmasking because CI users appear to be more sensitive to ILDs when either ILD or ITD cues are isolated and presented with fidelity, and because ILD cues are currently more readily available in clinical speech processors. As mentioned in the conclusion of van Hoesel *et al.* (2008), several details of the implementation of the PDT strategy may have adversely affected the utility of pulse timing for ITD sensitivity: (1) the use of 19 channels virtually ensures that current spread will produce channel cross-talk, blurring the presentation of ITDs from adjacent electrodes; (2) the use of channel-unique derivation of pulse timing and placing the lowest filter corner frequency at 250 Hz allows only a few of the most apical electrodes to carry ITDs at usable low rates; and (3) even usable (low-rate) ITD information is presented by unique pulse trains at each electrode, and therefore cues may be inconsistent, a serious confound if current spread is considered, in which case adjacent channels’ pulse timing may have provided conflicting or confusing cues to surviving auditory nerve populations in a region of current spread overlap. Hence, ITD information presented to listeners in that study may not have been carried under optimal conditions, and superior conveyance of ITDs may have provided more usable cues. Despite the apparent dominance of ILDs in CI spatial hearing, the addition of accessible and possibly redundant ITD information may provide the cues necessary for stream segregation and spatial release from masking (Ihfeldt and Litovsky, 2012).

The current results have shown stronger evidence for the binaural benefit of TFS-timed pulses than these earlier experiments. This may be due to differences in methodology or signal processing. First, the deleterious effects of spectral mismatch across ears due to electrode placement and current spread are thought to have been major factors in the observed reduced ability of bilateral CI listeners to utilize deliberately presented interaural TFS information (Poon *et al.*, 2009; Kan *et al.*, 2013). The present study attempted to minimize these effects by stimulating on pitch-matched electrode pairs. Furthermore, the use of only eight electrode pairs allowed for the physical separation of active electrodes, in order to reduce the possible effects of channel interaction due to current spread. Additionally, the present study avoided the potential confounds associated with the presentation of usable ITDs mentioned above by using low-rate (<200 Hz) pulse timing cues which were redundant across multiple electrodes. The data presented in Fig. 6 and discussed above were collected using a TFS strategy that was very similar to PDT in that each electrode's pulse timing was determined from the acoustic TFS in the corresponding channel. The similarity of results with these two strategies suggests that in each case, CI listeners were not benefiting from the added pulse timing information. The most likely impediments to using pulse timing ITDs were the presence of high rates in most of the channels and that the two channels carrying usable (<500 Hz) pulse timing information, the two most apical channels, were carrying the ITD cue at different rates. Overlap, due to current spread, between these two channels with different rates may possibly lead to less salient ITD cues compared to two channels with a common rate, because a common rate would provide some redundancy across channels, as well as a consistent inter-channel interference. The use of direct stimulation allows for excellent control of variables, and fills the gap between experiments using direct stimulation with modulated signals and free field experiments with clinical processors on a spectrum of realism. Among binaural CI direct stimulation experiments, the present one falls on the "realistic" side of that spectrum in the following ways. First, the CI signal processing is carried-out independently for left and right channels, and could theoretically be implemented on unlinked processors. Second, the use of speech stimuli for discrimination and lateralization reflects that speech is among the most commonly encountered acoustic signals of interest outside of the testing booth. Though its understanding can be achieved even when highly degraded, speech is a complex sound, with myriad temporal and spectral subtleties which contribute to its perceived qualities. It is highly encouraging that the TFS strategy used here could extract usable pulse timing information from the acoustic TFS of speech and that ITDs could be perceived even when imbedded in streams of dynamically changing rate. However, in order to study the effects of pulse timing on binaural hearing, we chose to introduce a rather large artificiality by including only ITD cues. As discussed above, the presence of ILD cues may overshadow the benefit from ITD cues in more realistic listening situations, and would certainly have complicated

the determination of pulse timing utility in these experiments.

The trade-off between better speech recognition at high rates and better pulse timing sensitivity at low rates presents a conundrum, and it may be instructive to discuss the underlying mechanisms. Envelope fluctuations may be better represented by high rate pulse trains due to two factors. First, a higher pulse rate represents a higher sampling rate, so that high frequency fluctuations will be more faithfully represented to the auditory nerve. The other factor is less obvious. In this study, the current levels at which stimuli are comfortable were found to be about 10–20 current units lower with high rates than with low rates. The drop in thresholds, though, was usually found to be greater, around 30–40 current units. Thus, the dynamic range for high rates can be 10–30 current units larger than for low rates, which is a 20% to 100% increase. This larger dynamic range may be responsible for superior envelope representation and speech understanding with high-rate stimulation. It is unknown which of these two phenomena dominates the improvement in speech understanding at high rates, but the answer may have implications for any implementation of low-rate stimulation for improved spatial hearing and suggests that additional techniques, such as pulse width modulation, may be needed to play a role in improving dynamic range for better envelope representation with low rates.

The current results indicate that CI listeners' spatial hearing improves without severe detriment to speech understanding in quiet when given binaural pulse timing cues at low and mixed rates. This finding suggests that some bilateral CI listeners may benefit from the inclusion of low- and/or mixed-rate TFS-derived pulse timing. However, it is unknown how the TFS and CIS strategies presented here would compare in realistic listening conditions, where listeners are also provided ILD and spectral tilt cues. Due to the independence of bilateral processing steps, we expect that the TFS strategy described here would provide the same ITD cues if it were implemented on independent bilateral processors. Additionally, although not directly tested here, the TFS strategy should have the ability to better represent  $f_0$  information, which could improve speech understanding in complex listening environments. Several research questions not addressed here are whether the localization abilities provided by this TFS strategy can produce spatial release from masking and whether melody recognition and music appreciation improve with this TFS strategy. These two important research avenues merit attention.

## V. CONCLUSION

The results from this experiment show that pulse timing can be a useful cue at low stimulation rates for ITD lateralization and discrimination of speech sounds with bilateral CIs. Pulse timing cues here derived from acoustic TFS were observed to be most useful when obtained from a low-frequency channel, resulting in low rates, and retained significant utility when presented on four apical electrodes in conjunction with high rates on four basal electrodes. Additionally, listeners generally understood speech in quiet

better when TFS-derived pulse timing cues were present, even though the TFS was not extracted from the same channels as the envelopes. Given the fact that these listeners are normally deprived of pulse timing cues representative of acoustic TFS, one would expect that their observed sensitivity to these cues might be lower here than if the listeners were routinely exposed to them. The tendencies observed here regarding differences in performance among listeners with different strategies and rate combinations suggest that future clinical devices should include processing strategies which preserve acoustic TFS information in pulse timing and perhaps have options for different rate combination settings.

## ACKNOWLEDGMENTS

The authors are grateful to the patients who participated in this research. This study was supported by NIH-NIDCD (5R01 DC003083, R.Y.L.; K99/R00 DC010206, M.J.G.) and also in part by a core grant to the Waisman Center from the NICHD (P30 HD03352).

- Algazi, V. R., Duda, R. O., and Thompson, D. M. (2001). "The CIPIC HRTF database," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York.
- Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W.-D., and Hamzavi, J.-S. (2007). "Speech and music perception with the new fine structure speech coding strategy: Preliminary results," *Acta Otolaryngol. (Stockh.)* **127**, 1298–1303.
- Bronkhorst, A., 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.
- Carlyon, R. P., Deeks, J. M., and McKay, C. M. (2010). "The upper limit of temporal pitch for cochlear-implant listeners: stimulus duration, conditioner pulses, and the number of electrodes stimulated," *J. Acoust. Soc. Am.* **127**, 1469–1478.
- Chatterjee, M., and Robert, M. E. (2001). "Noise enhances modulation sensitivity in cochlear implant listeners: stochastic resonance in a prosthetic sensory system?," *J. Assoc. Res. Otolaryngol.* **2**, 159–171.
- Churchill, T. H. (2014). "System and method for fine-structure processing for hearing assistance devices," U.S. Patent No. 8,694,114 (April 8, 2014).
- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* **110**, 1150–1163.
- Galvin, J. J., 3rd, and Fu, Q. J. (2005). "Effects of stimulation rate, mode and level on modulation detection by cochlear implant users," *J. Assoc. Res. Otolaryngol.* **6**, 269–279.
- Goldsworthy, R. L., and Shannon, R. V. (2014). "Training improves cochlear implant rate discrimination on a psychophysical task," *J. Acoust. Soc. Am.* **135**, 334–341.
- 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.
- Ihlefeld, A., and Litovsky, R. Y. (2012). "Interaural level differences do not suffice for restoring spatial release from masking in simulated cochlear implant listening," *PLoS ONE* **7**, e45296.
- 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.
- Kiang, N. Y., and Moxon, E. C. (1972). "Physiological considerations in artificial stimulation of the inner ear," *Ann. Otol.* **81**, 714–730.
- Kidd, G., Jr., Best, V., and Mason, C. R. (2008). "Listening to every other word: Examining the strength of linkage variables in forming streams of speech," *J. Acoust. Soc. Am.* **124**, 3793–3802.
- Kidd, G., Mason, C. R., Best, V., and Marrone, N. (2010). "Stimulus factors influencing spatial release from speech-on-speech masking," *J. Acoust. Soc. Am.* **128**, 1965–1978.
- Laback, B., and Majdak, P. (2008). "Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates," *Proc. Natl. Acad. Sci.* **2007**, 814–817.
- 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.
- Litvak, L. M., Delgutte, B., and Eddington, D. K. (2001). "Auditory nerve fiber responses to electric stimulation: Modulated and unmodulated pulse trains," *J. Acoust. Soc. Am.* **110**, 368–379.
- 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.
- Loizou, P. C., Poroy, O., and Dorman, M. (2000). "The effect of parametric variations of cochlear implant processors on speech understanding," *J. Acoust. Soc. Am.* **108**, 790–802.
- 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**(4), 352–360.
- 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**(3), 1479–1490.
- Macpherson, E. A., and Middlebrooks, J. C. (2002). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," *J. Acoust. Soc. Am.* **111**, 2219–2236.
- Majdak, P., Laback, B., and Baumgartner, W.-D. (2006). "Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing," *J. Acoust. Soc. Am.* **120**, 2190–2201.
- Misurelli, S. M., and Litovsky, R. Y. (2012). "Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry," *J. Acoust. Soc. Am.* **132**, 380–391.
- Moore, B. C. J. (2003). "Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants," *Otol. Neurotol.* **24**, 243–254.
- Nie, K., Stickney, G., and Zeng, F.-G. (2005). "Encoding frequency modulation to improve cochlear implant performance in noise," *IEEE Trans. Biomed. Eng.* **52**, 64–73.
- 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**, 806–905.
- Riss, D., Arnoldner, C., Reiss, S., Baumgartner, W.-D., and Hamzavi, J.-S. (2009). "1-year results using the Opus speech processor with the fine structure speech coding strategy," *Acta Otolaryngol. (Stockh.)* **129**, 988–991.
- Rubinstein, J. T., and Miller, C. A. (1999). "How do cochlear prostheses work?," *Curr. Opin. Neurobiol.* **9**, 399–404.
- Rubinstein, J. T. Y., Wilson, B. S., Finley, C. C., and Abbas, P. J. (1999). "Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation," *Hear. Res.* **127**, 108–118.
- Schatzer, R., Krenmayr, A., Au, D. K. K., Kals, M., and Zierhofer, C. (2010). "Temporal fine structure in cochlear implants: Preliminary speech perception results in Cantonese-speaking implant users," *Acta Otolaryngol. (Stockh.)* **130**, 1031–1039.
- Schleich, P., Nopp, P., and D'Haese, P. (2004). "Head shadow, squelch, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant," *Ear Hear.* **25**, 197–204.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Sit, J.-J., Simonson, A. M., Oxenham, A. J., Faltys, M. A., and Sarpeshkar, R. (2007). "A low-power asynchronous interleaved sampling algorithm

- for cochlear implants that encodes envelope and phase information," *IEEE Trans. Biomed. Eng.* **54**, 138–149.
- Smith, R. (1979). "Adaptation, saturation, and physiological masking in single auditory-nerve fibers," *J. Acoust. Soc. Am.* **65**, 166–178.
- Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). "Chimaeric sounds reveal dichotomies in auditory perception," *Nature* **416**, 87–90.
- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech Hear. Res.* **28**, 455–462.
- van Hoesel, R. J. M. (2004). "Peak-derived timing stimulation strategy for a multi-channel cochlear implant," U.S. patent 2,004,172,101.
- van Hoesel, R. J. M., Böhm, 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," *J. Acoust. Soc. Am.* **123**, 2249–2263.
- van Hoesel, R. J. M., and Clark, G. M. (1997). "Psychophysical studies with two binaural cochlear implant subjects," *J. Acoust. Soc. Am.* **102**, 495–507.
- 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.
- van Hoesel, R. J. M., and Tyler, R. S. (2003). "Speech perception, localization, and lateralization with bilateral cochlear implants," *J. Acoust. Soc. Am.* **113**, 1617–1630.
- Vermeire, K., Punte, K., and Heyning, V. D. (2010). "Better speech recognition in noise with the fine structure processing coding strategy," *ORL J. Otorhinolaryngol. Relat. Spec.* **72**, 305–311.
- Wichman, F., and Hill, N. (2001). "The psychometric function: I. Fitting, sampling, and goodness of fit," *Percept. Psychophys.* **63**, 1293–1313.
- Wightman, F., and Kistler, D. (1992). "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Am.* **91**, 1648–1661.
- Wilson, B., Finley, C., Lawson, D., Wolford, R., Eddington, D. K., and Rabinowitz, W. (1991). "Better speech recognition with cochlear implants," *Nature* **352**, 236–238.
- Zierhofer, C. (2003). "Electrical nerve stimulation based on channel specific sampling sequences," U.S. Patent No. 2,003,105,504.