M1

# Mixed Stimulation Rates for Encoding Temporal Cues in Bilateral Cochlear Implants


Tyler Churchill[1], Alan Kan[1], Matt Goupell[2], Ruth Litovsky[1]

[1]University of Wisconsin-Madison; [2]University of Maryland

*e-mail: tchurchill@wisc.edu*


Binaural Hearing and Speech Laboratory


*Conference on Implantable Auditory Prostheses*

*Lake Tahoe, CA*

*July 14th – 19th, 2013*


## INTRODUCTION

Bilateral cochlear implants (CIs) provide clear benefits to recipients, but there is a significant gap in performance on spatial hearing tasks relative to performance of normal-hearing (NH) listeners. Much of the ability of NH listeners to localize sounds in the horizontal plane depends upon being able to extract and utilize the interaural time differences (ITDs) in the "temporal fine structure" (TFS) of sound between the two ears at frequencies below 1.5 kHz (Wightman and Kistler 1992; Macpherson and Middlebrooks 2002). Most contemporary CIs stimulate the auditory nerve with constant-rate, high-rate electric pulse trains and discard the TFS (Wilson et al. 1991; Arndt et al. 1999). These strategies may allow for adequate envelope ITD representation, but the loss of TFS may account for many of the deficits in binaural hearing abilities compared to NH. Though CI listeners can use the timing cues in low-rate pulse trains for ITD discrimination and monaural rate discrimination, better speech recognition is generally achieved with high-rate stimulation (Loizou et al. 2000; Galvin and Fu 2005). The present study explored this tradeoff between better speech recognition with high rates and pulse timing usability for ITDs with low rates.

## METHODS

Three tasks were used to systematically examine the effects of channel pulse rate and pulse timing on ITD discrimination, ITD lateralization, and recognition of speech in quiet by CI listeners. Stimuli consisted of speech tokens processed at low, high, and mixed electrical stimulation rates, presented synchronously on 8 binaurally pitch-matched electrode pairs. For each rate combination, two strategies were tested: a novel strategy which preserves a representation of acoustic TFS information in the pulse timing and the Continuous Interleaved Sampling (CIS) strategy, which does not.

**Speech recognition in quiet (100 words per condition)**

•4 or 5-word sentences (from Kidd et al. 2008 corpus) were presented from 0º azimuth in quiet (ITD = 0 μs)

**ITD L/R discrimination (40 reps per condition)**

•ITDs (±1600, ±800, ±400, ±200, ±100, ±50 μs) were applied to sentence corpus names (e.g., "Gene", "Mike", etc.) and presented in a 2I2AFC task (L→R or R→L)

**ITD lateralization (10 reps per condition)**

•ITDs calculated for azimuths from -70º to +70º in 10º increments were applied to corpus names

•Listener's task was to indicate the perceived source location along a 180º arc via graphical user interface

**Rate combinations**

1. low rates (100-173 Hz) on all electrodes
2. mixed rates; low rates on four apical electrode pairs, high rates on four basal electrode pairs
3. high rates (894-1547 Hz) on all electrodes

**Processing Strategies**

1. CIS: constant-rate pulse train carriers with average rates equal to average rates of corresponding stimulus token from TFS strategy
2. TFS: pulse train carrier consists of pulses timed to the positive-going zero-crossings of $\mathcal{Im}(\mathcal{H}(y))$, where $\mathcal{Im}$ denotes taking the imaginary part, $\mathcal{H}$ denotes the Hilbert transform, and y is the band-pass filtered acoustic signal appropriate for low-rate (100-173 Hz) or high-rate (894-1547 Hz) temporal fine structure extraction

Envelopes for both strategies were extracted via a short-time Fourier transform method using 512-point Blackman windows with 256-point overlaps and compressed into the listeners' dynamic ranges using a power law function with exponent = 1/3. Frequency bins were allocated into 8 channels with corner frequencies logarithmically-spaced between 195 and 16015 Hz.

## RESULTS

### Listeners

| Subject Code | Age | Hearing Aid Use (years) | Left/Right CI Use (years) |
|---|---|---|---|
| IAJ | 68 | 46 | 17/9 |
| IBF | 51 | 14 | 5/7 |
| IBK | 73 | 8 | 10/4 |
| IBM | 59 | 13 | 3/7 |
| IBN | 66 | 50 | 3/13 |
| IBR | 59 | 22 | 5/9 |
| ICD | 56 | 40 | 4/10 |
| ICM | 60 | 29 | 3/1 |
| AVG: from pooled listener responses | | | |

**Table 1.** Eight bilateral CI listeners participated in these experiments. All subjects were Cochlear devices and used the ACE processing strategy.

### Speech Recognition

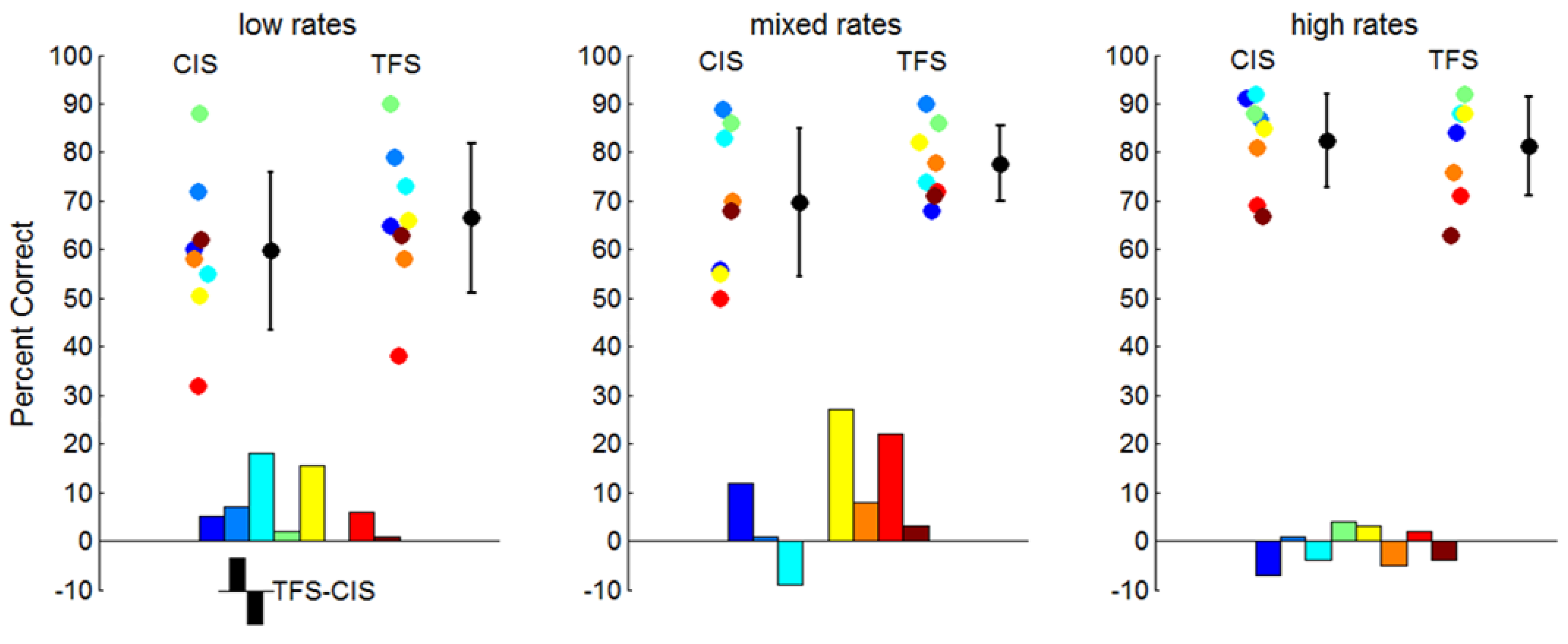


**Figure 1.** Speech recognition in quiet. Subjects' percent correct scores are shown above for each strategy and rate combination. Error bars on mean markers indicate standard deviation of listener scores. Colored bars at the bottom of each plot indicate the differences between listeners' average scores with TFS and CIS strategies for each rate combination.

### Discrimination vs Lateralization

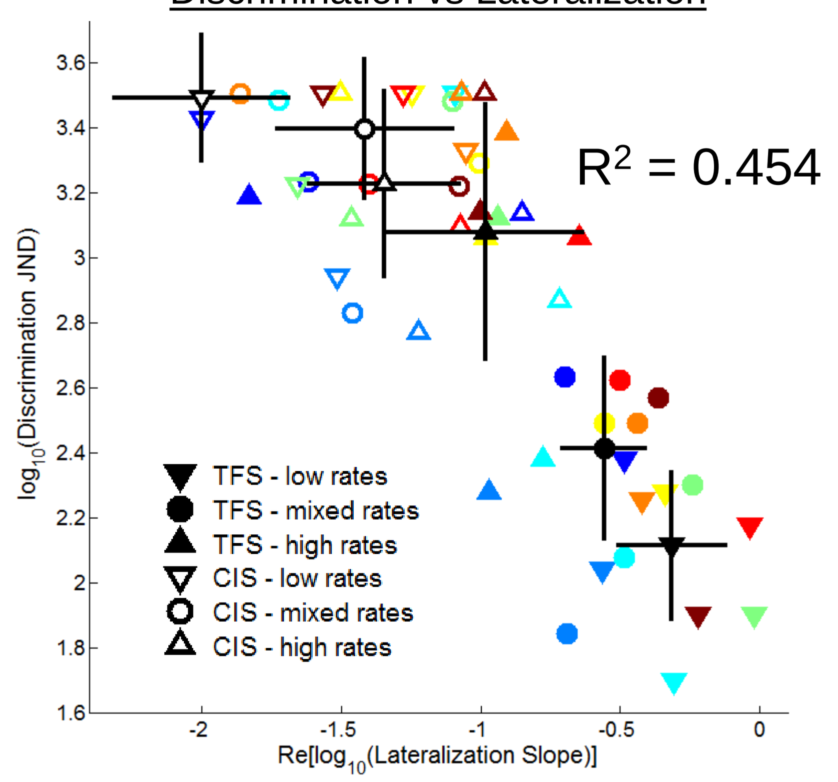


**Figure 4.** Comparison of lateralization slopes and discrimination JNDs reveals a correlation of spatial hearing measures across conditions.

### Pairwise t-tests

Arcsin-Transformed % Correct Speech Recognition

| | | CIS LOW | CIS MIXED | CIS HIGH | TFS LOW | TFS MIXED |
|---|---|---|---|---|---|---|
| CIS | MIXED | 0.30 | | | | |
| CIS | HIGH | **0.043** | 0.30 | | | |
| TFS | LOW | 0.18 | 0.81 | 0.062 | | |
| TFS | MIXED | 0.07 | 0.55 | 0.81 | 0.30 | |
| TFS | HIGH | **0.043** | 0.30 | 0.81 | **0.049** | 0.81 |

Lateralization Slopes

| | | CIS LOW | CIS MIXED | CIS HIGH | TFS LOW | TFS MIXED |
|---|---|---|---|---|---|---|
| CIS | MIXED | 0.78 | | | | |
| CIS | HIGH | 0.78 | 0.78 | | | |
| TFS | LOW | **0.0052** | **0.0056** | **0.0165** | | |
| TFS | MIXED | **0.0013** | **0.0011** | **0.0212** | 0.08 | |
| TFS | HIGH | **0.018** | 0.07 | 0.57 | **0.0128** | **0.0165** |

Discrimination JNDs

| | | CIS LOW | CIS MIXED | CIS HIGH | TFS LOW | TFS MIXED |
|---|---|---|---|---|---|---|
| CIS | MIXED | 0.71 | | | | |
| CIS | HIGH | 0.45 | 0.71 | | | |
| TFS | LOW | **0.0016** | **0.0057** | **0.0378** | | |
| TFS | MIXED | **0.0015** | **0.0084** | **0.04** | **0.0468** | |
| TFS | HIGH | 0.055 | 0.097 | 0.21 | **0.0378** | **0.0419** |

**Table 2.** Significance levels ($P < 0.05$ in **bold**) of Holm-corrected pairwise t-tests of arcsin-transformed percent correct speech recognition scores, lateralization slopes and discrimination JNDs. A binary logistic regression performed on speech recognition responses found significant effects of rate combination ($P < 0.001$) and the interaction of strategy and rate combination ($P = 0.012$), but not of strategy alone ($P = 0.516$).

### Lateralization

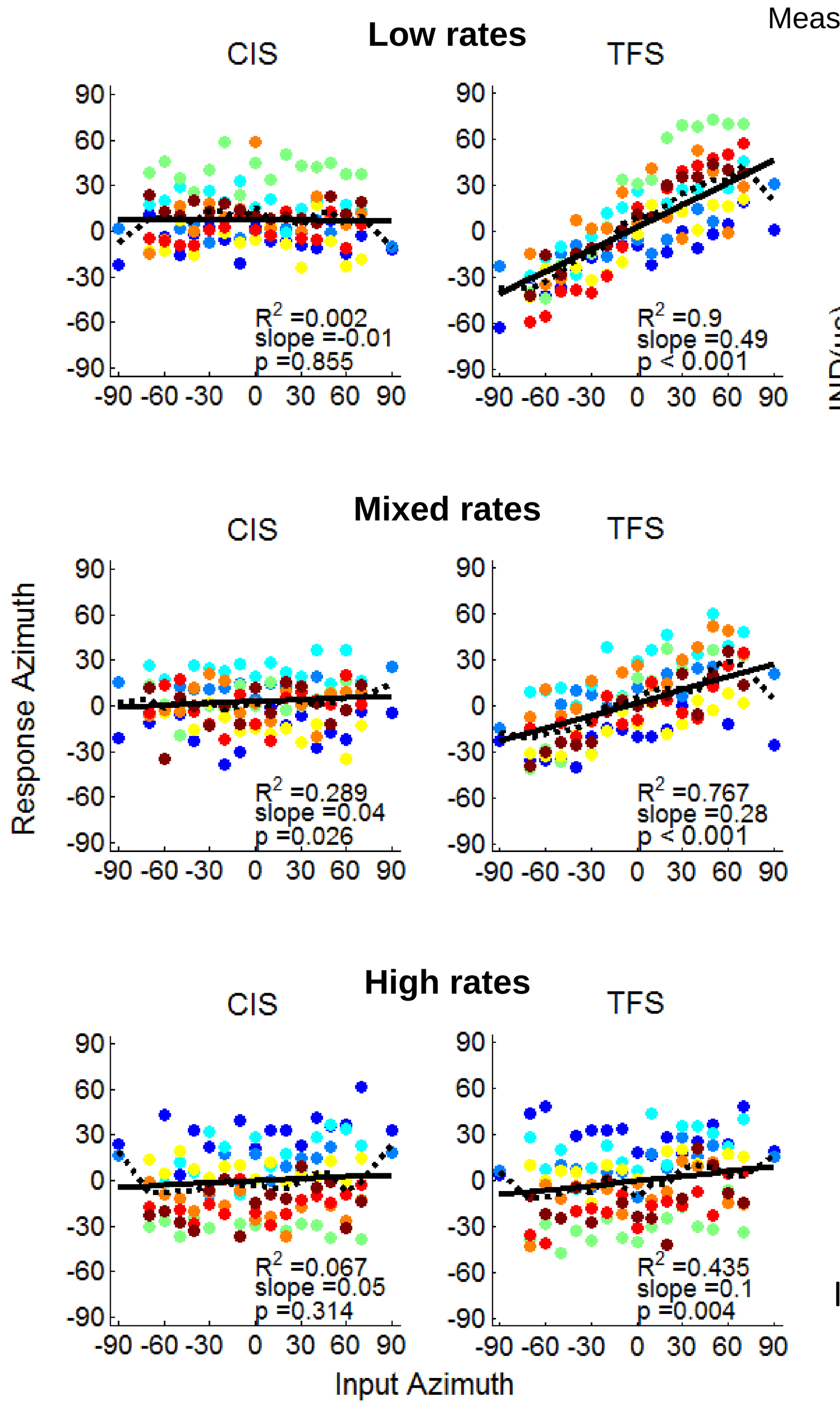


**Figure 2.** Listeners' average lateralization responses to speech tokens containing only ITD location cues. Dashed lines indicate averages across listeners, and solid lines indicate linear best fits, for which statistics are included in the lower-right corner of each plot.

### Discrimination

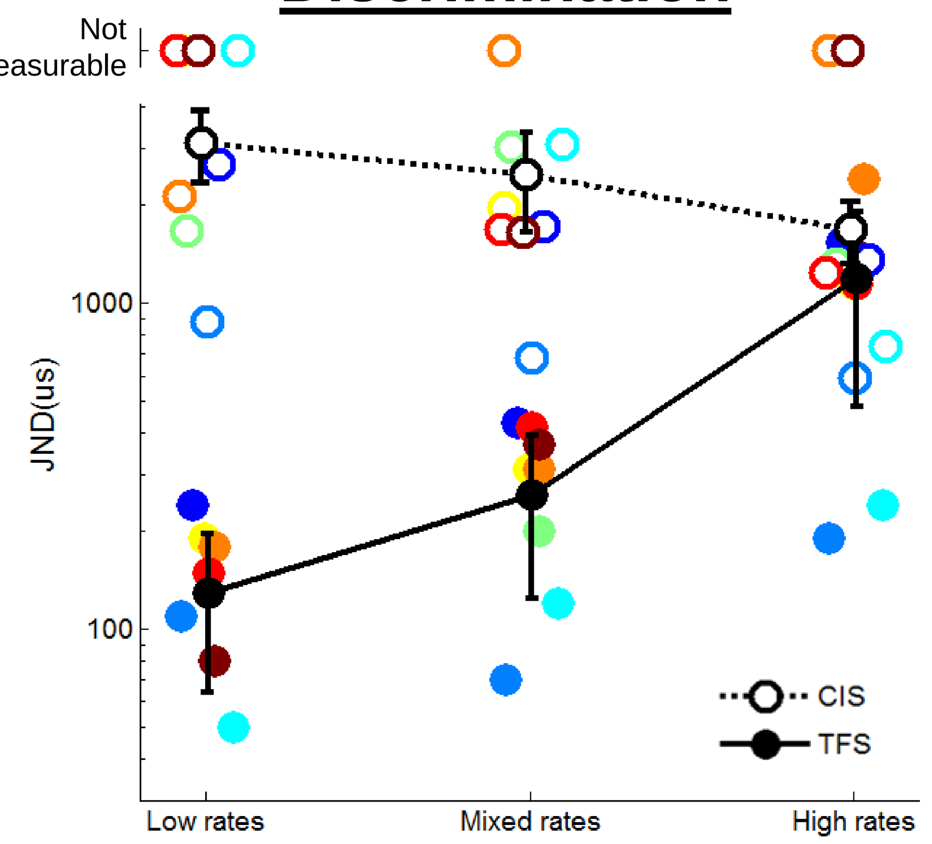


**Figure 3** Just-Noticeable Differences (JNDs) for L/R discrimination. "AVG" JNDs were calculated from pooled listener data, and error bars indicate standard deviation of listeners' measurable JNDs.

### Analysis of Variance

| | Recognition | Lateralization | Discrimination |
|---|---|---|---|
| Strategy | NS | ***P* < 0.001** | ***P* < 0.001** |
| Rate | ***P* < 0.001** | ***P* < 0.001** | NS* |
| Interaction | NS | ***P* < 0.001** | ***P* = 0.002** |

**Table 3.** Significance levels of strategy, rate, and rate-strategy interaction effects due to ANOVA in arcsin-transformed % correct speech recognition, lateralization slopes, and discrimination JNDs. *Within-subjects (repeated measures) ANOVA in discrimination JNDs, accounting for variability at each set of conditions, also revealed a significant effect of rate combination ($P < 0.001$).

### Discrimination vs Speech Recognition

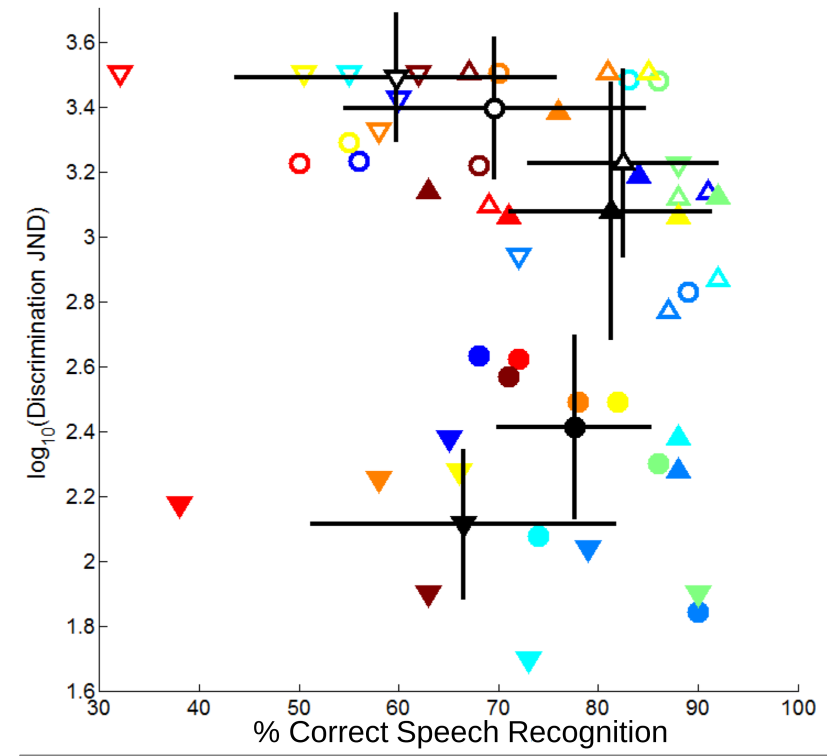


**Figure 5.** Comparison of discrimination JNDs and speech recognition scores reveals the speech/spatial performance tradeoff.

### Lateralization vs Speech Recognition

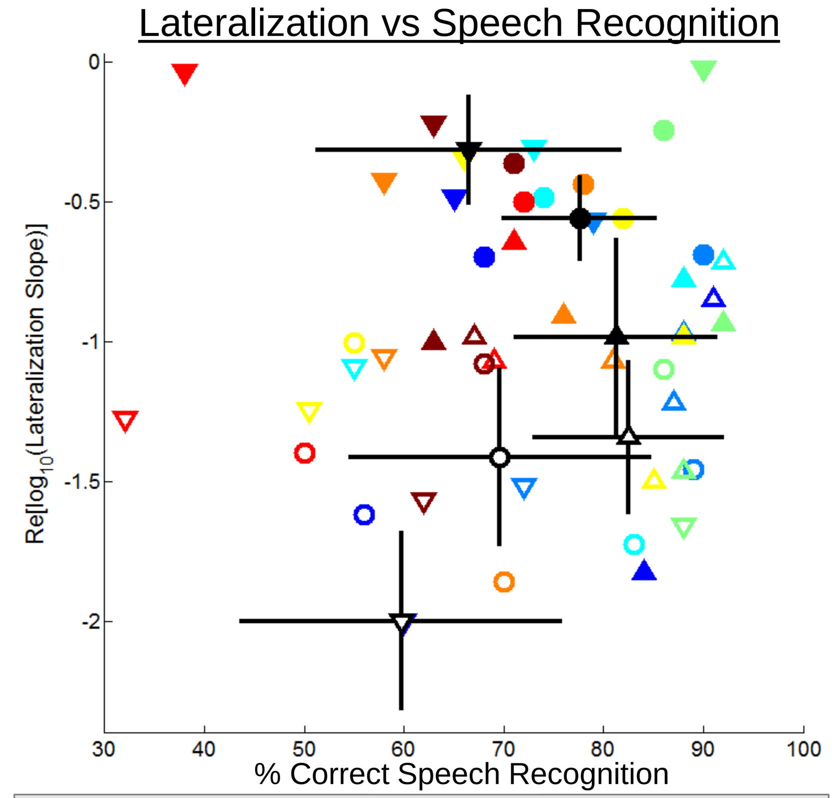


**Figure 6.** Comparison of lateralization slopes and speech recognition scores reiterates the tradeoff as a function of pulse rate.

## DISCUSSION

This study has quantitatively characterized the tradeoff between binaural pulse timing sensitivity at low rates and speech recognition at high rates using test and control processing strategies that were largely unfamiliar to the listener. The convergence of average discrimination JNDs and lateralization slopes for both strategies at high rates as seen in Figures 2 and 3 reflects the decreasing accessibility of pulse timing information and the increasing reliance on envelope cues, which were available with both strategies. As shown in Figure 1, superior speech recognition in quiet with both strategies at high rates suggests better envelope representation.

Although the ability to lateralize and discriminate ITDs may rely on different mechanisms, Figure 4 suggests a relationship between the measures. An ANOVA in discrimination JNDs as a function of corresponding lateralization slopes reveals a significant relationship [$F(1,46) = 39.3$, $P < 0.001$].

The ability of listeners to receive the benefits of high and low rates when presented with both (mixed rate condition) suggests that a mixed-rate paradigm may provide an optimal cue set for speech recognition and binaural timing sensitivity. Indeed, some listeners exhibited both good speech recognition and ITD sensitivity with low- and mixed-rate stimuli, as indicated by data points in the lower-right corner of Figure 5 and the upper-right corner of Figure 6.

Because ITDs were applied to acoustic signals prior to processing for CIs, and because left and right channels were subsequently processed independently, current results suggest that implementation of TFS-representing strategies on unlinked bilateral processors could provide accurate binaural timing cues. Due to the importance of these cues for such phenomena as spatial release from masking, further research is warranted.

## CONCLUSIONS

- Bilateral CI listeners were able to use low-rate and mixed-rate pulse timing cues derived from acoustic TFS in speech tokens for ITD discrimination and lateralization of multi-channel stimuli.
- Speech recognition was degraded with low-rate stimuli relative to mixed- and high-rate stimuli.
- Bilateral CI listeners could benefit from processing strategies which included TFS-timed pulses at low and/or mixed stimulation rates.

## REFERENCES


Arndt, P, Staller, S, Arcaroli, J, Hines, A and Ebinger, K (1999). *Within-subject comparison of advanced coding strategies in the Nucleus 24 cochlear implant* (Cochlear Corporation Report for FDA submission). April, 1999

Galvin, J J, and Fu, Q J (2005). "Effects of stimulation rate, mode and level on modulation detection by cochlear implant users." J Assoc Res Otolaryngol **6**(3): 269-279.

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**(6): 3793-3802.

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.

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**(5): 2219.

Wightman, F and Kistler, D (1992). "The dominant role of low-frequency interaural time differences in sound localization." J Acoust Soc Am **91**(3): 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).


## ACKNOWLEDGEMENTS


Work supported by NIH-NIDCD (5R01 DC003083, Litovsky) and a core grant to the Waisman Center from the NICHD (P30 HD03352).