PS-175

WAISMAN CENTER

Neural coding of acoustic temporal fine structure and envelope: Psychophysiological assessment of peripheral encoding on sound localization

Heath Jones*, Tyler Churchill, Alan Kan, and Ruth Y. Litovsky
University of Wisconsin–Madison
e-mail: hgjones@wisc.edu

Binaural Hearing and Speech Laboratory

37th Midwinter ARO Meeting
San Diego, CA
February 22-26th, 2014

INTRODUCTION

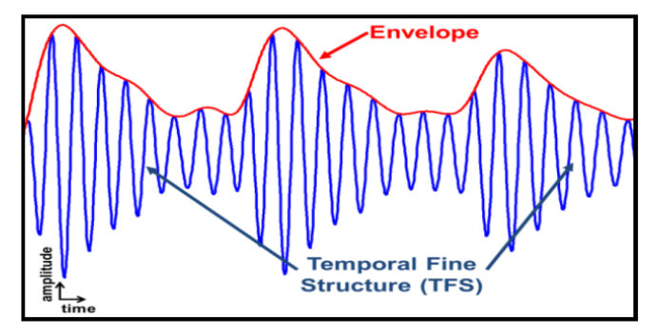
- One of the major difficulties in improving the poor sound localization abilities of bilateral cochlear implant (CI) patients is that we do not understand well the transformations of acoustical signals due to peripheral encoding across the large population of auditory nerve (AN) fibers in normal hearing (NH) listeners.

?
Acoustics → **Neural Processing** → **Behavior**

- Most of our understanding comes from either single neuron recordings[1] or psychoacoustical lateralization/discrimination studies that use carefully constructed amplitude-modulated stimuli.
- Here we combine simulated neural responses across a large population of ANs spanning the audible spectrum to "virtual free-field speech" stimuli with psychoacoustical sound localization measures using the same stimuli.

The present study aims to bridge the gap between single unit physiology and psychoacoustical measures of behavior

- Complex acoustic waveforms, such as speech, can be decomposed into a slowly varying "envelope" (ENV) and a rapidly-varying "temporal fine-structure" (TFS).



- Recently, several studies have investigated the importance of ENV and TFS coding by the auditory periphery for speech perception[2,3] using computational AN models[5,6].
- However, the roles of neural ENV and TFS coding to the "binaural processing" of sound localization cues in NH listeners has not been addressed.

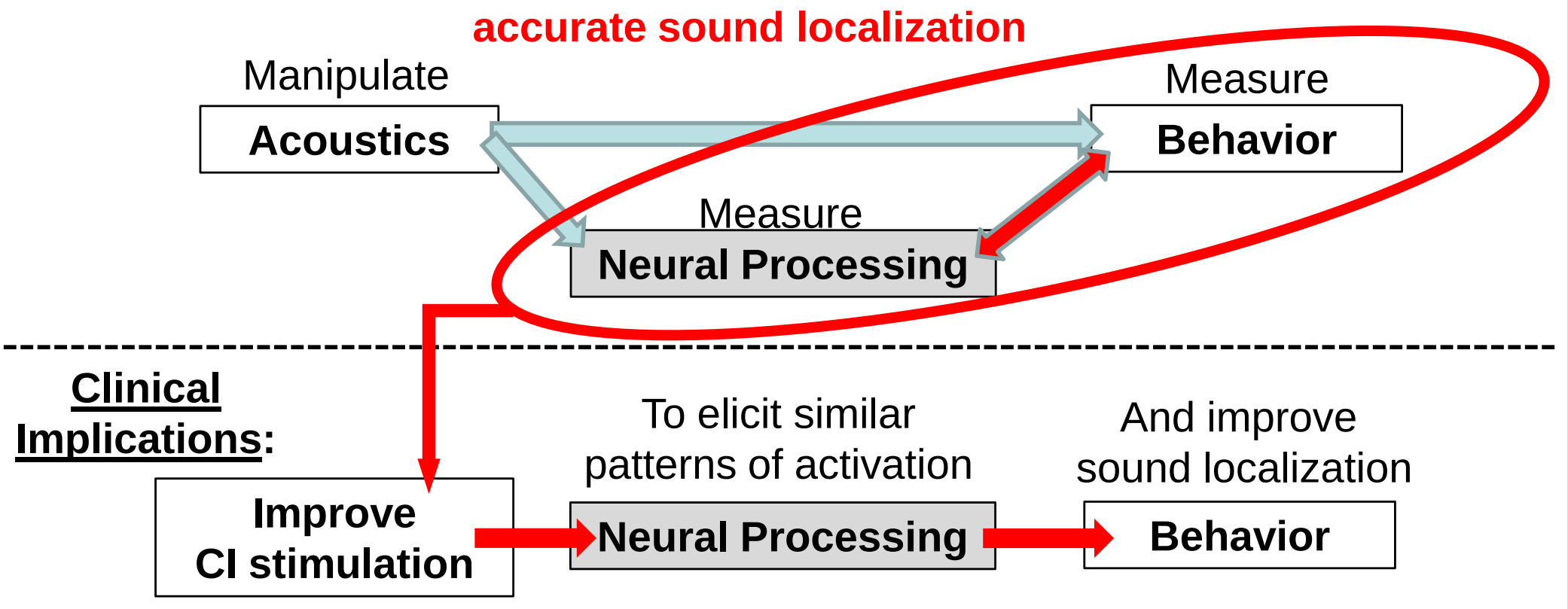
Outputs of a computational AN model[5,6] for "virtual free-field speech" were analyzed across the ears and binaural metrics were extracted from measures of neural TFS and ENV coding

These binaural metrics were then used to predict psychophysical localization responses

"THE QUICK TOUR"

(Numbers and colors correspond to poster panels)

Goal of study: To gain insight into the broad patterns of neural activation responsible for accurate sound localization



1.) A sound localization experiment conducted with NH listeners showed that vocoded "virtual free-field speech" envelopes do not produce accurate localization, regardless of carrier.

2.) Using the localization experiment stimuli, a binaural computational model analysis showed unique patterns of activation as a function of azimuthal location for neural TFS, but neural ENV exhibited no consistent patterns.

3.) Outputs of the neural population binaural model predicts human sound localization results.

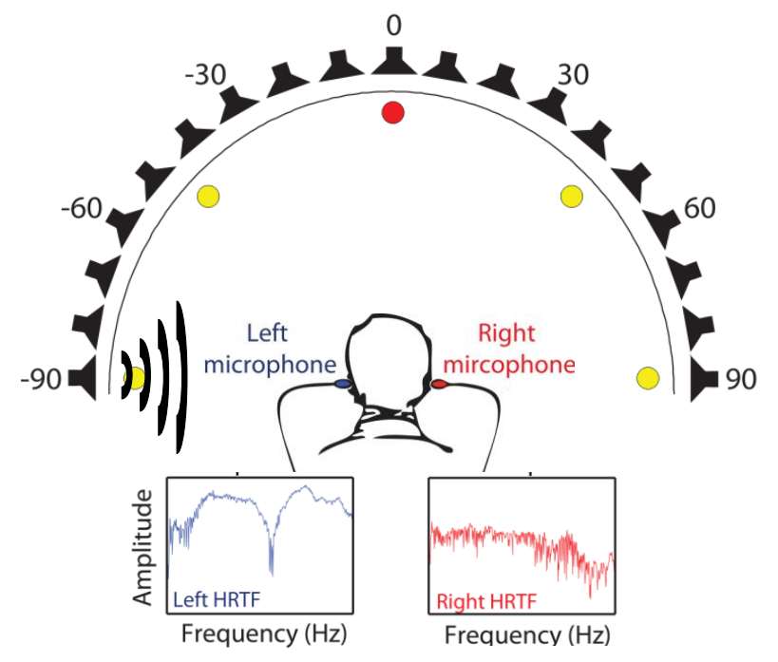
4.) The binaural computational model presented here can be used to inform development of CI speech processing strategies that promote improved sound localization abilities.

VIRTUAL FREE-FIELD SPEECH

HRTF Measurements

- Blocked-meatus microphone pair at ear canal entrances (AuSIM)
- Golay codes (5 reps, 200ms duration)

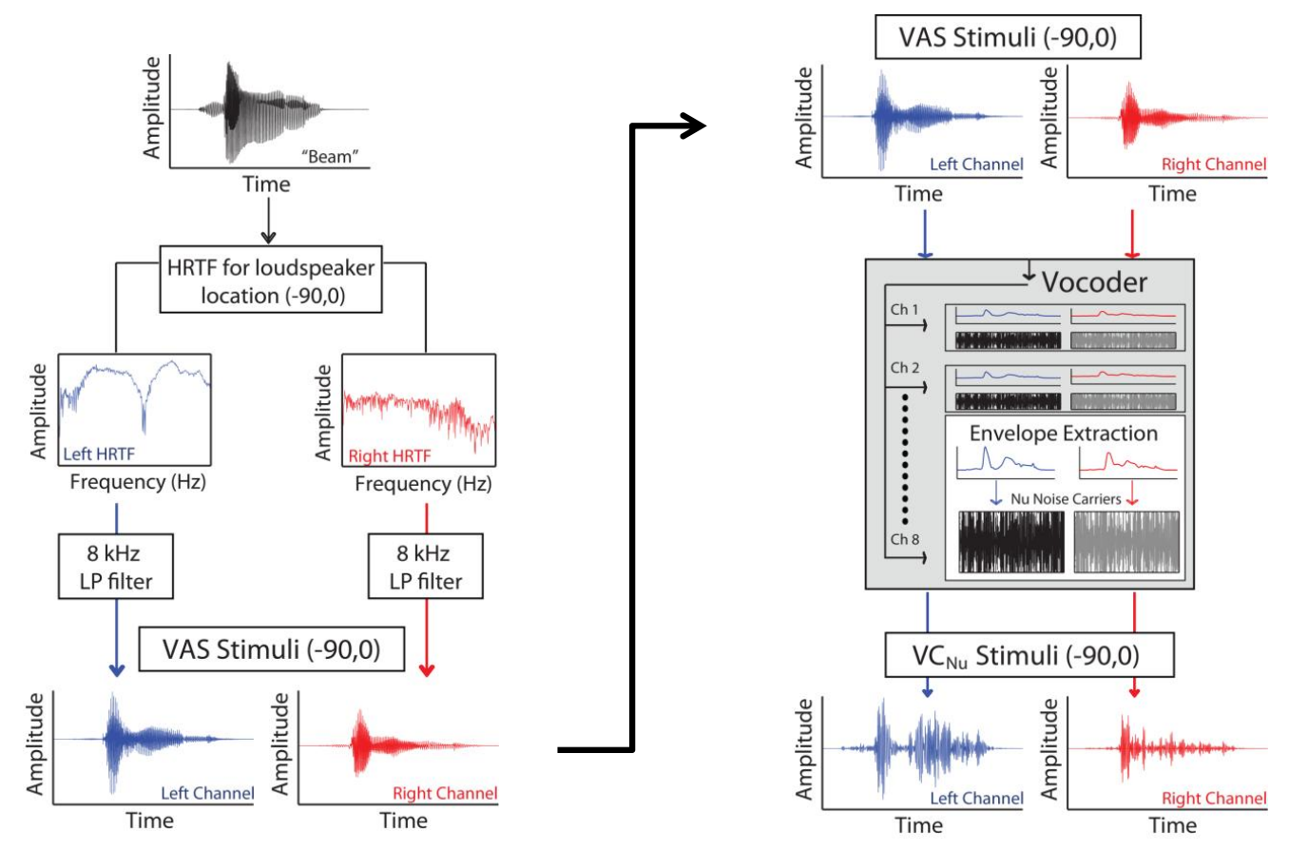
Measurement of head-related transfer functions (HRTFs)



Stimuli

- Five consonant-nucleus-consonant (CNC) male spoken words

Virtual Acoustic Space (VAS) stimuli | **Vocoded (VC) stimuli**



Stimulus Generation Schematic. (Left) Virtual acoustic space (VAS) stimuli were generated for 19 loudspeaker locations using individualized HRTFs measured for each listener. **(Right)** VAS stimuli were generated then processed by an 8-channel noise vocoder with uncorrelated noise carriers (VC_{Nu}).

Use stimuli to measure sound localization performance

Input stimuli into AN model

1.) LOCALIZATION PERFORMANCE

Listeners

- 8 normal hearing adults (ages 20-30)

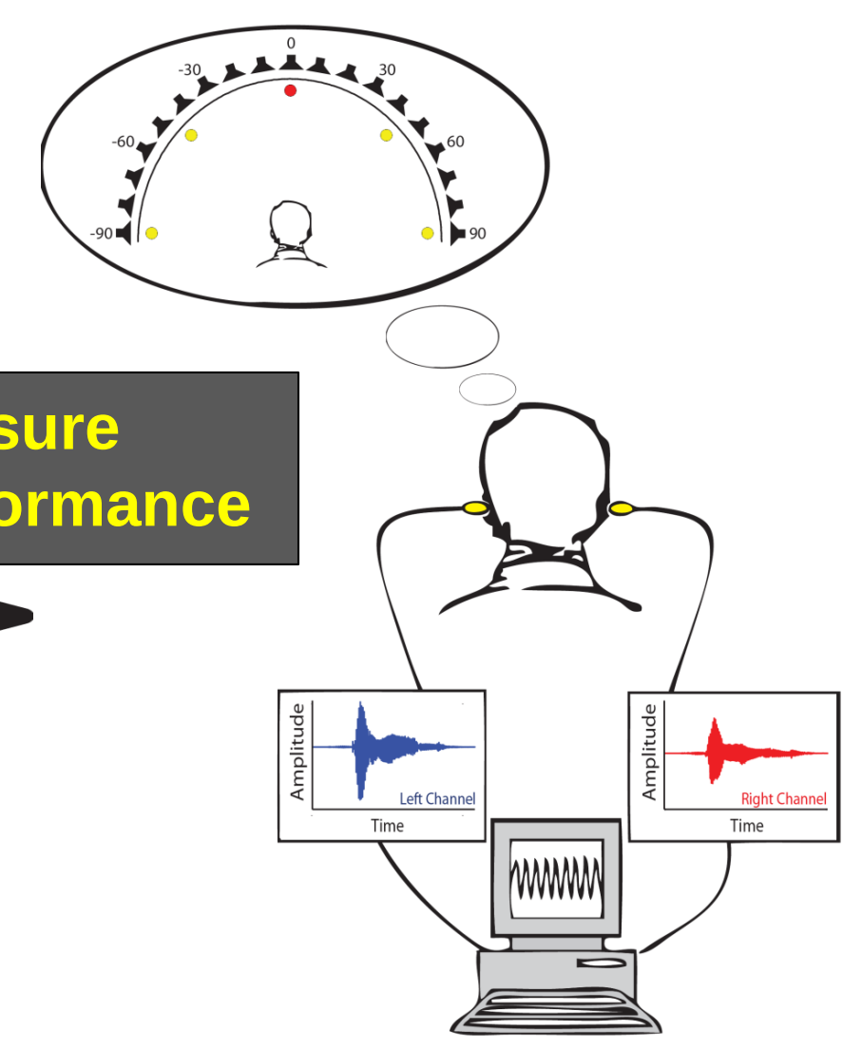
Stimuli

Same five (CNC) male spoken words

- Virtual acoustic space (VAS)
- Uncorrelated noise (VC_{Nu})

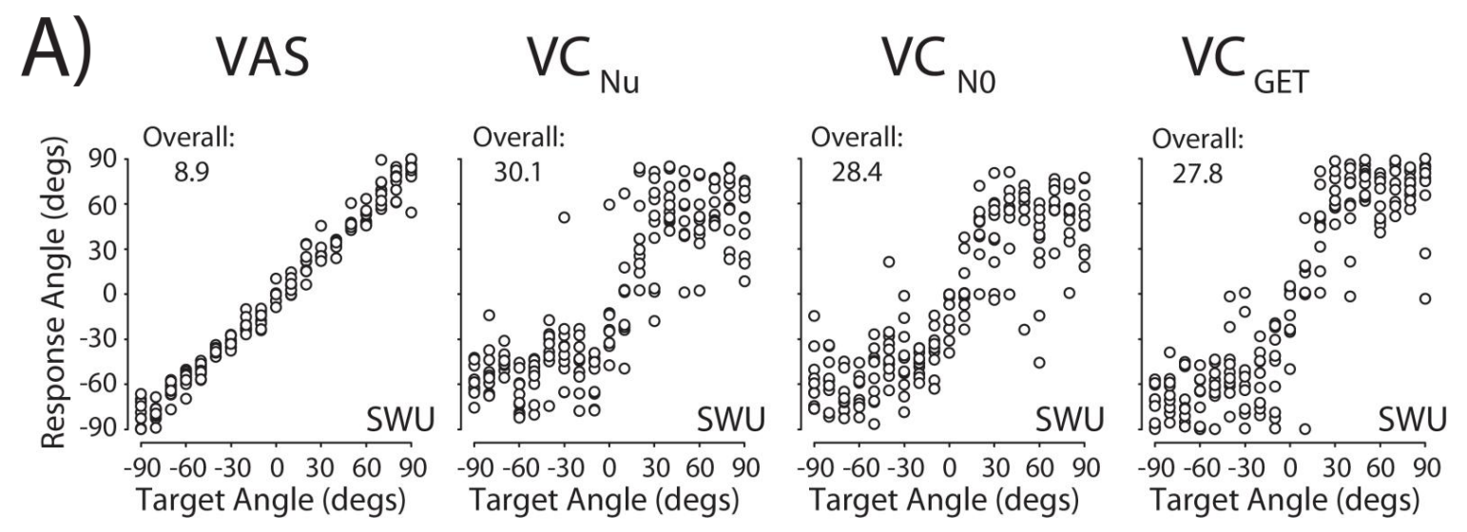
Additional stimuli

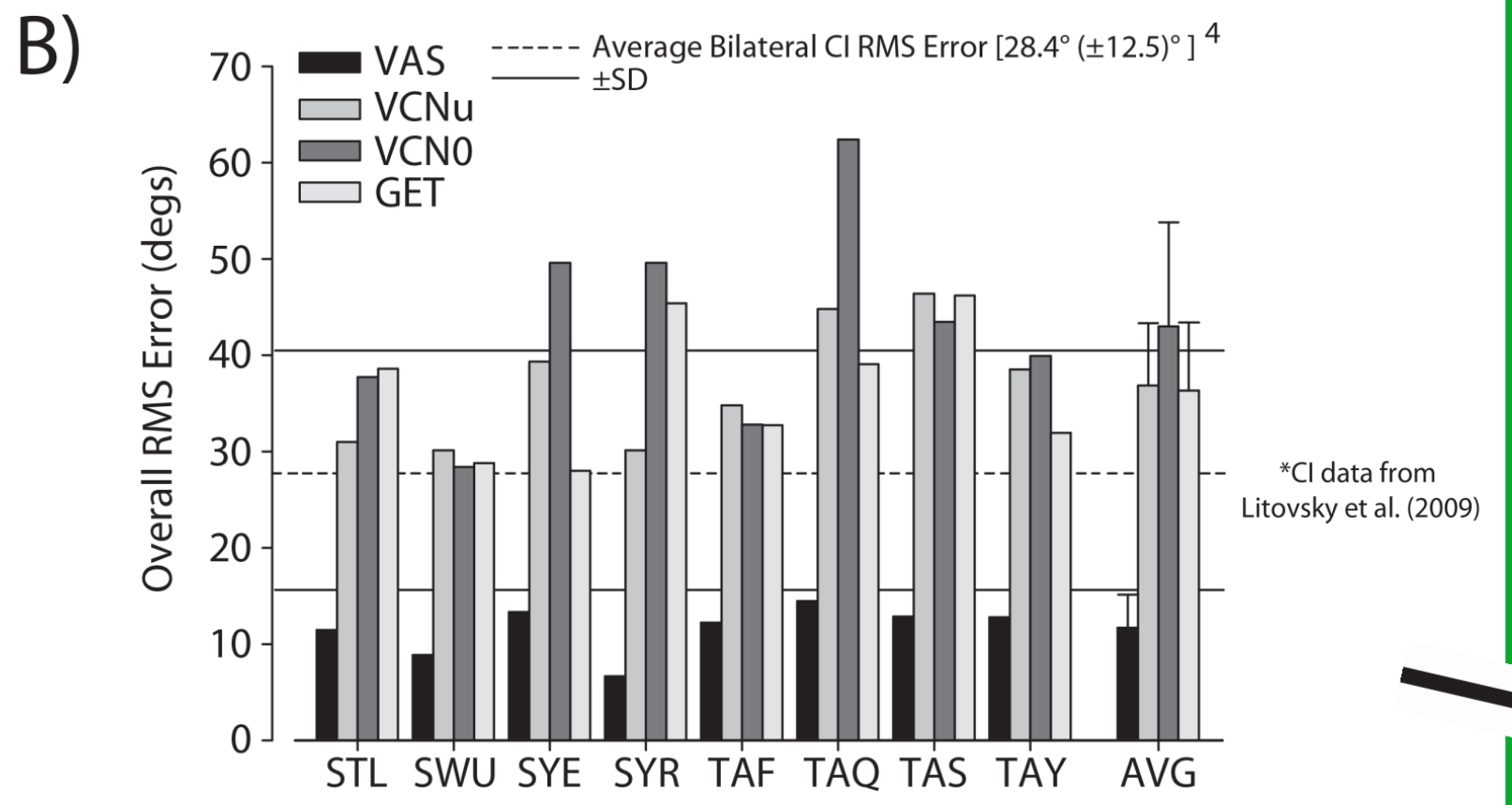
- Correlated noise (VC_{N0})
- Gaussian envelope tones (VC_{GET})



Procedure

- Stimuli presented twice from each location (2 x 5 words = 10 reps) at 60dB SPL (roved ±4dB SPL)
- Subjects indicated on computer screen the perceived source location





Synchronized carrier stimulation did not improve performance compared to uncorrelated carriers

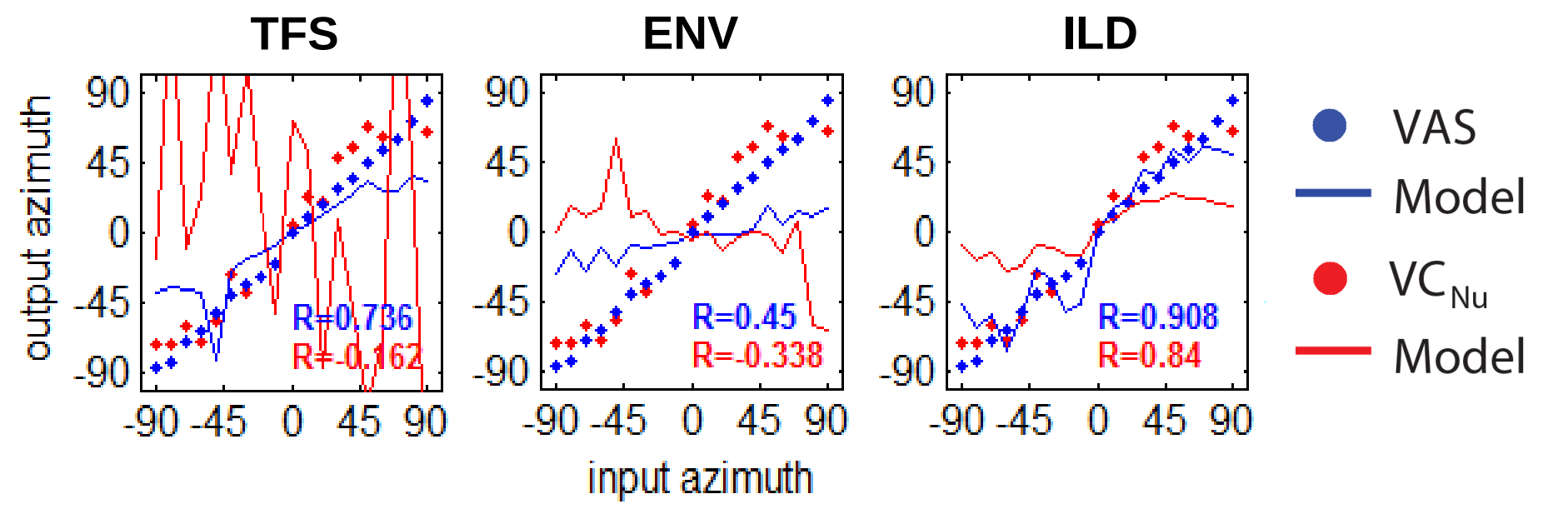
3.) MODEL PREDICTIONS

Procedure

- Model coefficients (β) were calculated from least-squares fits of binaural metrics and response data for VAS stimuli
- Only the ILD metric could partially predict the listener's response for vocoded stimuli

Binaural metrics calculated from neural spike outputs were able to predict psychophysical localization responses

$$RMS\ Error = \beta \times \delta_b + constant$$



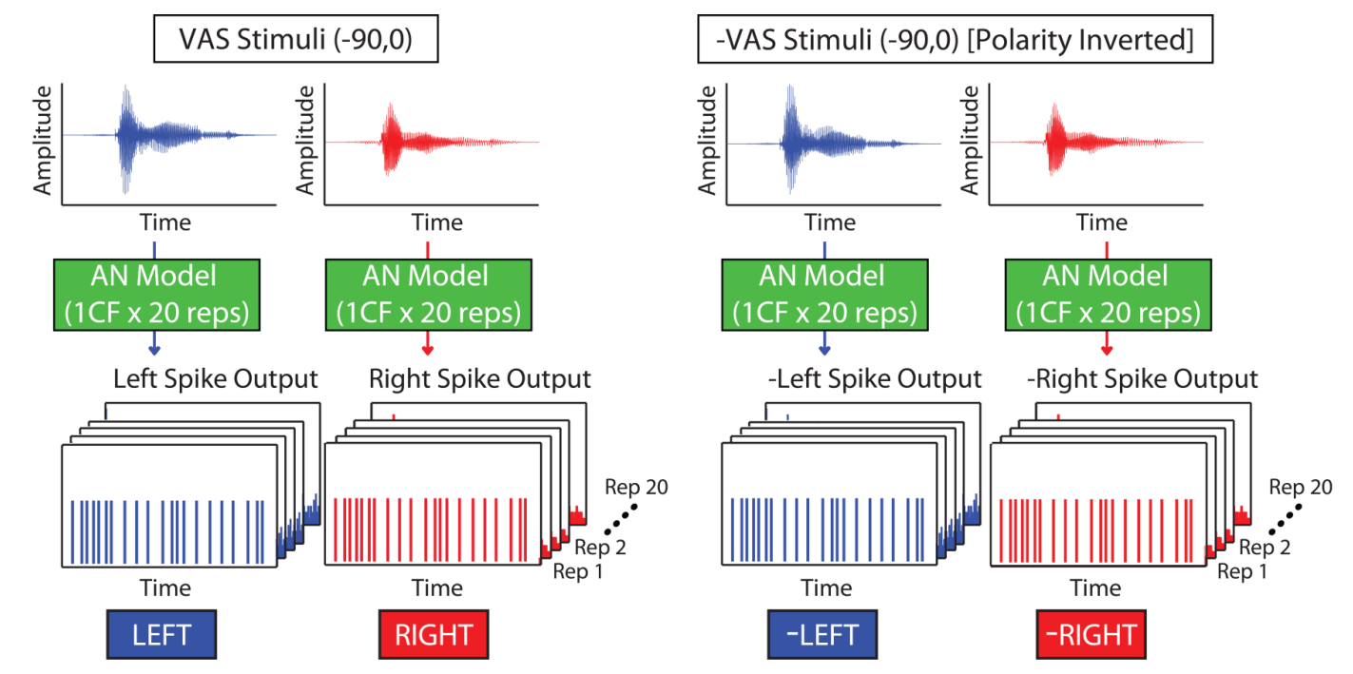
Model responses and correlation coefficients for VAS and VC_{Nu} stimuli. Average listener responses (dots) and model (blue or red line) predictions for each virtual azimuth. The far right panel shows a linear model of ILD prediction of VC_{Nu} responses generated for those stimuli.

Combine 1.) sound localization measures and metrics of 2.) neural population binaural model

2.) NEURAL POPULATION BINAURAL MODEL

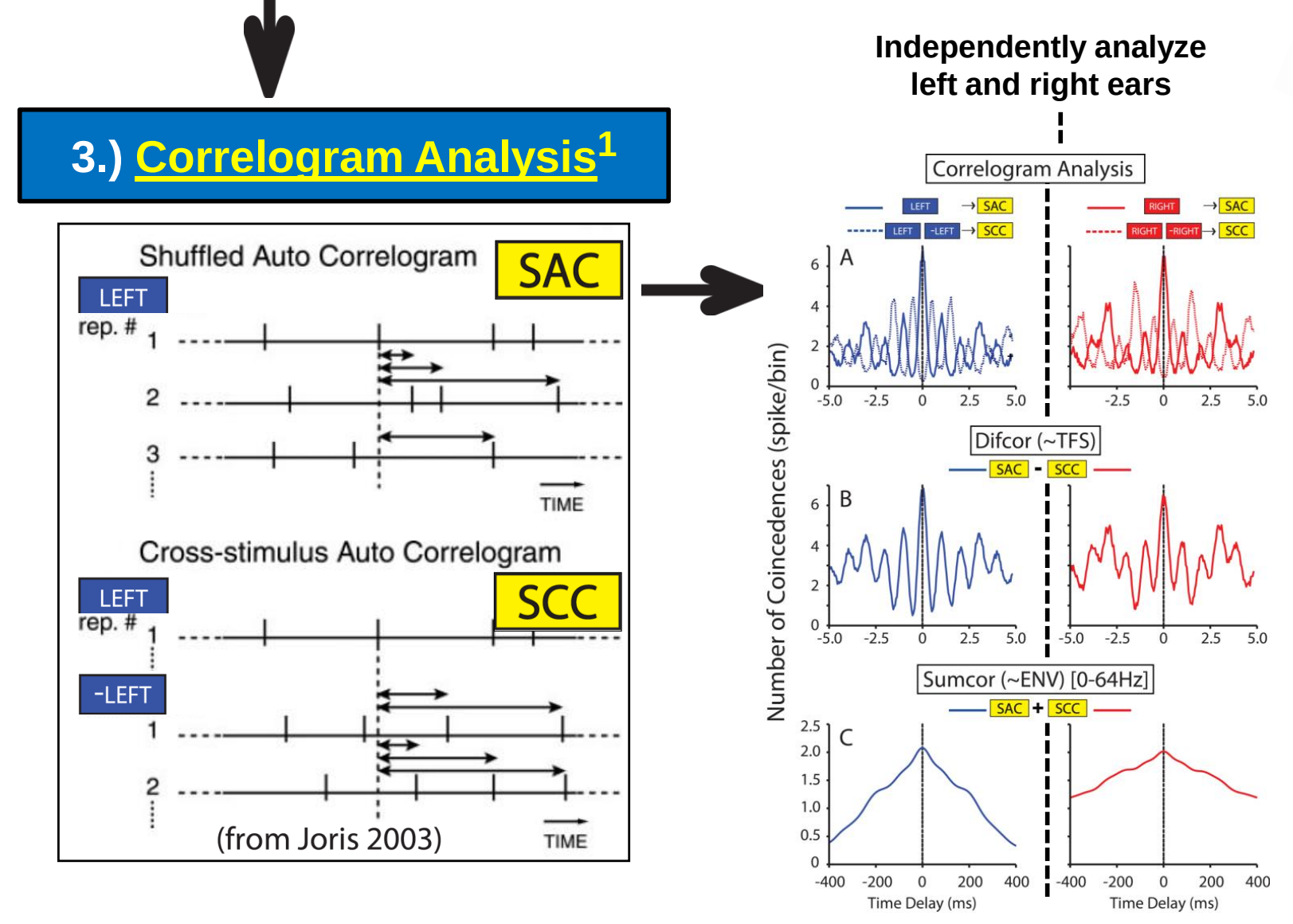
Neural ENV and **neural TFS** refer to temporal aspects of action potential firing which are **unaffected** or largely **affected** by a polarity-inversion of the input stimulus, respectively and can be evaluated using "Correlogram Analysis"[1]

1.) Input original and inverted stimuli independently into AN Model[5,6] and obtain neural spike outputs (20 reps)

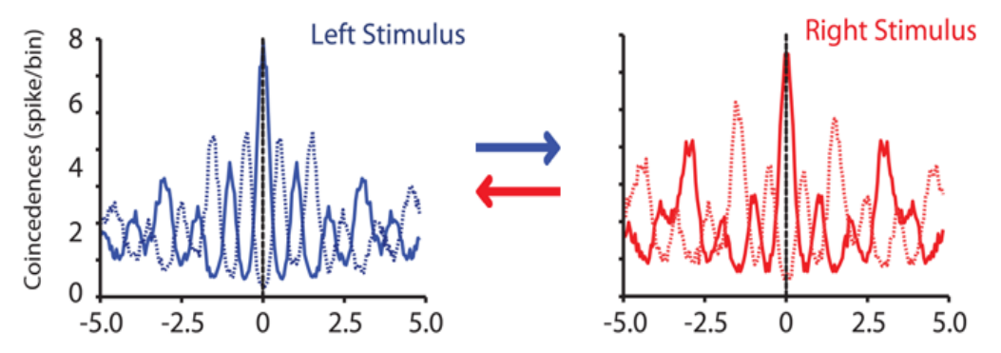


2.) 100 fibers with characteristic frequencies (CFs) spaced logarithmically (0.1 - 16 kHz) were simulated

3.) Correlogram Analysis[1]

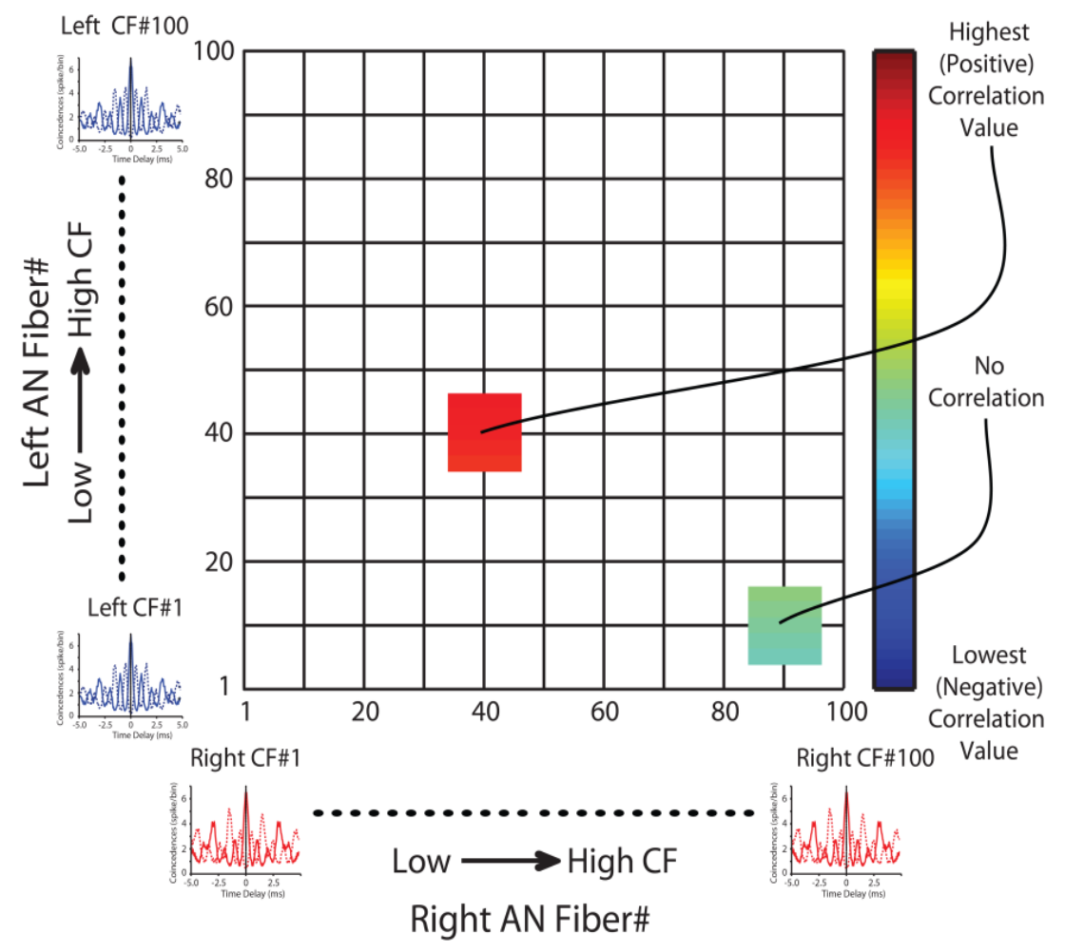


Analyze outputs between the ears



4.) Correlogram analysis of neural spike between ears across all CFs

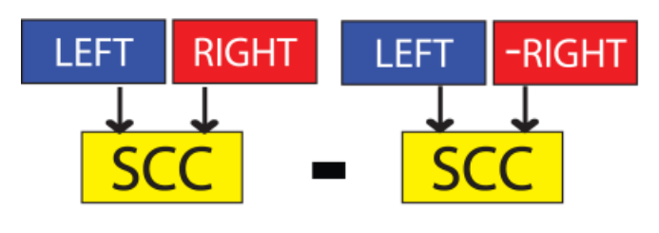
BINAURAL CORRELATION PLANE (BCP)

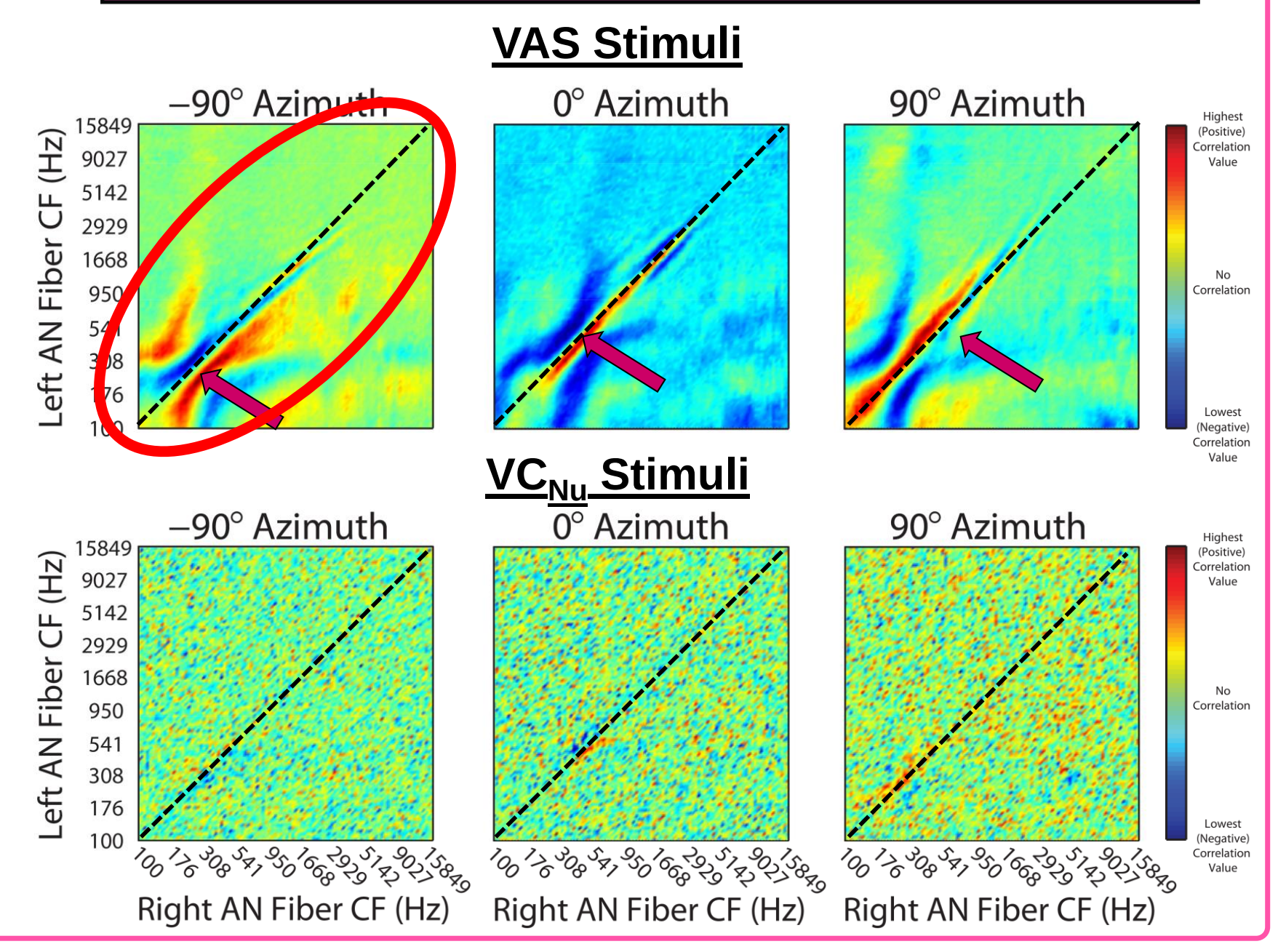


5.) Binaural correlograms values at zero time delay were calculated for each of the 100 left ×100 right CF pairs

BINAURAL CORRELATION PLANE (BCP)

Binaural Difcor (TFS) =

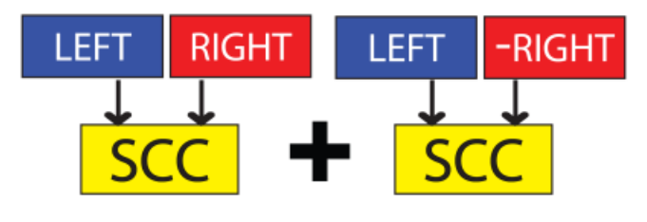


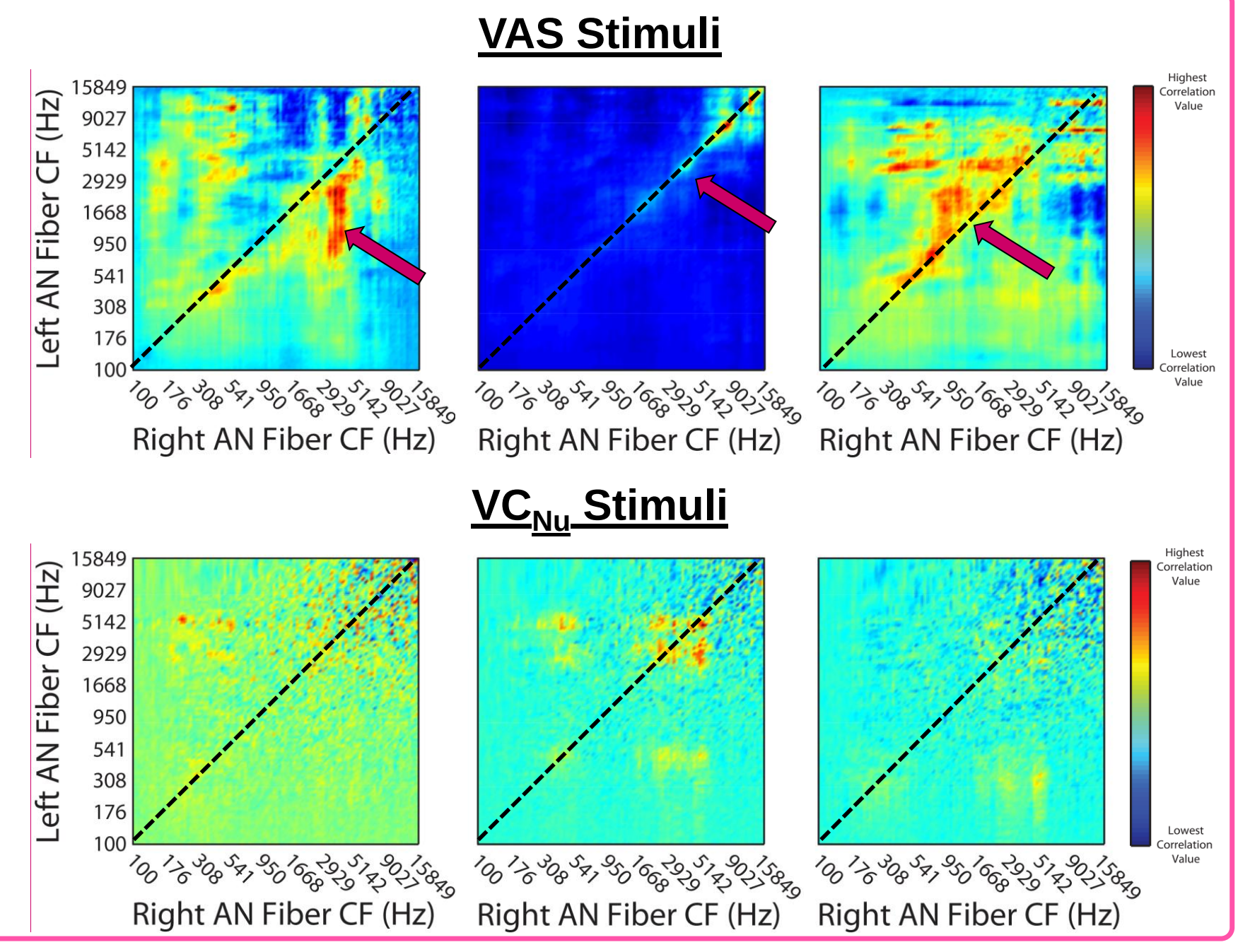


Interaural neural TFS revealed unique activation patterns for each location (red arrow), but not for vocoded stimuli

Spike outputs between the ears showed high correlations for off-setting CFs at all non-zero source locations[4]

Binaural Sumcor (ENV) =



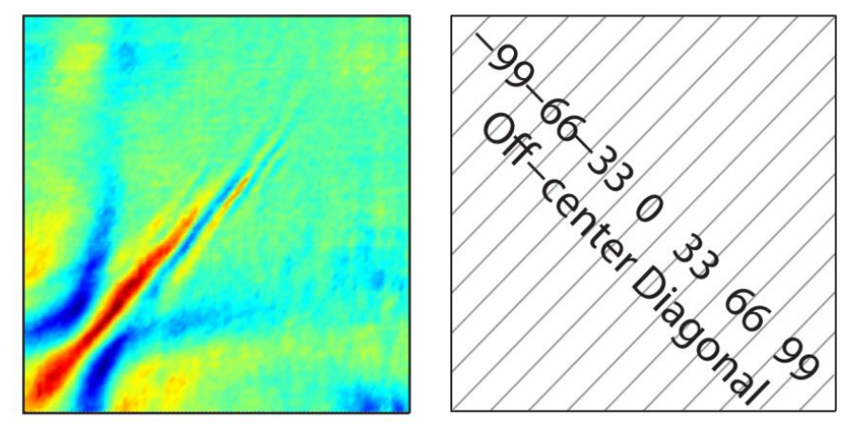


Neural ENV correlations exhibited a smaller area of activation both across frequency and location

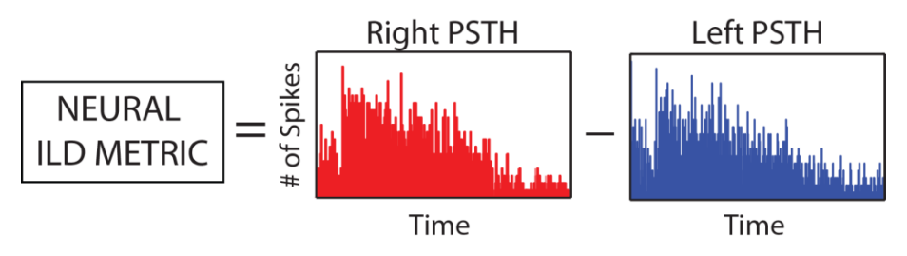
Unique patterns of activation along the BCP were not apparent for neural ENV coding for vocoded stimuli

BINAURAL METRICS

BCP values summed along the diagonal axes and normalized by the number of fiber comparisons



The abscissae of the peaks in the BCP diagonal sums (δ_b) were extracted as metrics and were found to be generally dependent upon generating azimuth for neural TFS.



A Neural ILD Metric was calculated by subtracting the left total spike count from the right total spike count

4.) CONCLUSIONS

- **Psychophysical localization responses could be predicted from simulated binaural metrics obtained from the outputs of a neural computational model.**
- Interaural correlations in spike timings exhibited a "stereausis processing"[7] pattern of activation, suggesting that ITDs could be encoded by an array of coincidence detection cells innervated by inputs with slightly off-setting CFs.
- The synchronizing of pulsatile stimulation (i.e., VC_{GET} vocoded stimuli) across the ears did not result in improved sound localization performance compared to other vocoder carriers.
- The neural population binaural model presented here gives us the ability to explore new stimulation strategies and provides insight to the psychophysical performance that could be expected.

REFERENCES

[1] Joris, P. X. (2003). Interaural time sensitivity dominated by cochlea-induced envelope patterns. *J Neurosci*, *23*(15), 6345–6350.

[2] Shamma, S. A., & Lorenzi, C. (2013). On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system. *J Acoust Soc Am*, *133* (5), 2818–2833.

[3] Swaminathan, J., & Heinz, M. G. (2012). Psychophysiological Analyses Demonstrate the Importance of Neural Envelope Coding for Speech Perception in Noise, *J Neurosci*, *32* (5), 1747–1756.

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

[5] Zilany, M. S. A. and Bruce, I. C. (2006). "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," *J Acoust Soc Am, 120* (3):1446–1466.

[6] Zilany, M.S.A., Bruce, I.C., Nelson, P.C., and Carney, L.H. (2009). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: Long-term adaptation with power-law dynamics," *J Acoust Soc Am, 126* (5), 2390-2412.

[7] Shamma, S. A., Shen, N., and Gopalaswamy, P. (1989). "Stereausis: Binaural processing without neural delays. *J Acoust Soc Am*, *66*(3), 989–1006.

ACKNOWLEDGEMENTS

We would like to thank all our listeners. This work is supported by NIH-NIDCD (R01 DC003083, R. Litovsky) and in part by a core grant to the Waisman Center from the NICHD (P30 HD03352).