The Computational Sensorimotor Systems Laboratory

The magnetoencephalographic spectrotemporal response function in auditory cortex

Francisco Cervantes Constantino[1], Jonathan Z. Simon[2,3,4]

[1]Program in Neuroscience and Cognitive Science, [2]Department of Electrical & Computer Engineering, [3]Department of Biology, [4]Institute for Systems Research
University of Maryland, College Park

Introduction

The spectro-temporal receptive field (STRF), based on spike responses, is a classic tool in auditory electrophysiology used for quantitative and qualitative analysis of cortical function.

The analogous spectro-temporal *response function* derived from MEG responses can be used in the same way.

MEG-based STRF models reveal considerable predictive power and consistency, regardless of the stimulus used to generate them – from artificial multitone patterns to natural sounds.

Methods

Stimuli. 50 s auditory scenes were composed of pseudo-random temporally fixed tones from a pool of 10 frequency values (range: 180-2144 Hz), interspaced by 2 ERB steps. Five presentation rates (2, 4, 6, 8, or 10 per second) of tone clouds were used. Within each frequency channel, tone onsets were uniformly distributed with minimum inter-tone gaps of 40 ms.

Experiment design. Each of the 5 main scenes was present in 4 different blocks, for a total of 20 trials, with a concomitant task to ensure listener's attention. Scenes were interleaved with other similar but random scenes of varying tone rates, and participants (N=15) were asked to report rate changes via a button press. Trial order was randomized, and durations ranged between 70 and 120 s.

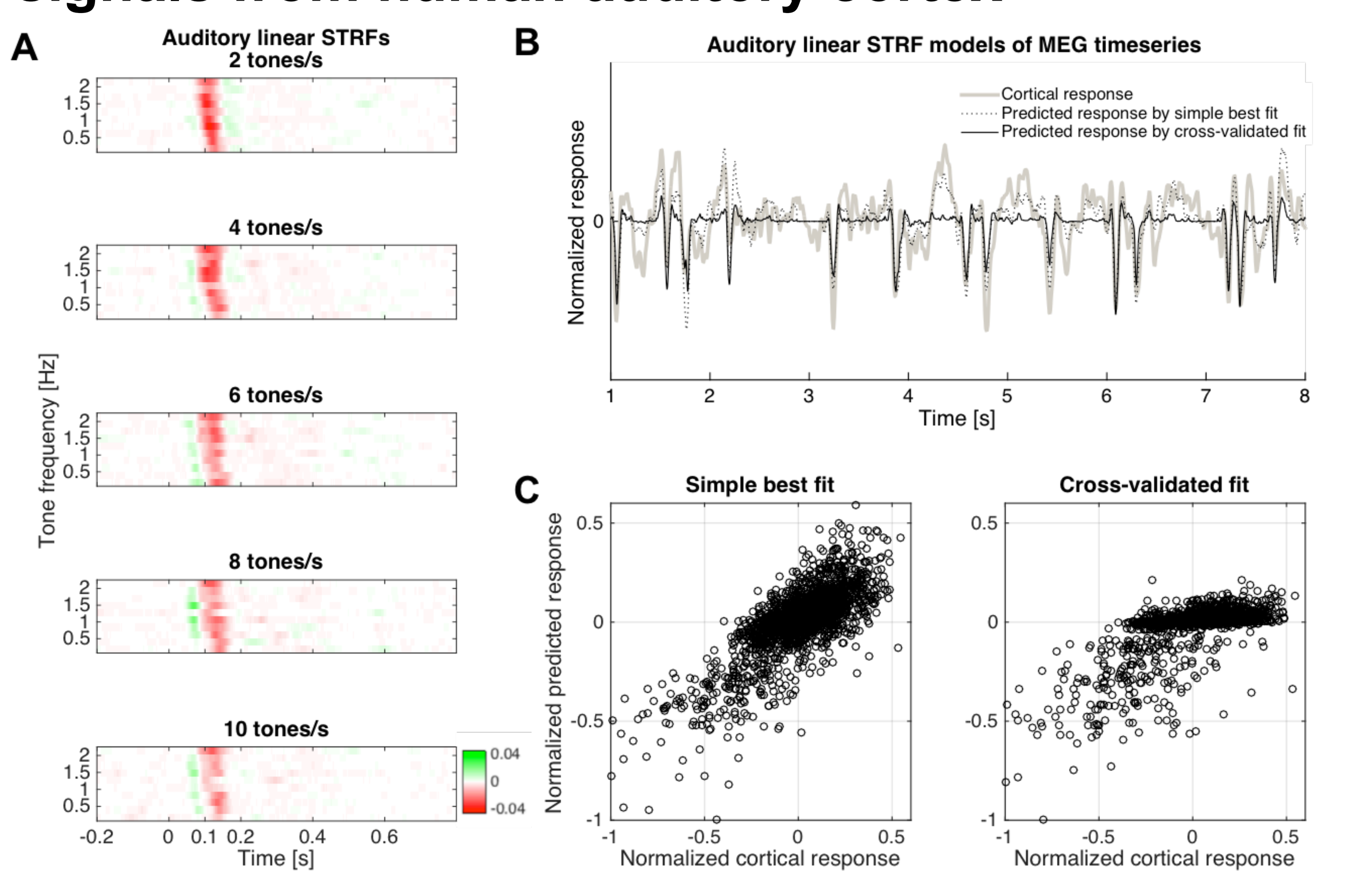
Data acquisition and analysis. Environmental and sensor noise contributions to neural signals from a 157-channel, whole-head MEG-KIT system (1 KHz sampling rate and 60 Hz notch filter) were estimated and removed. Sensor recordings were band-pass filtered 1-15 Hz then spatially filtered into a single virtual sensor data reflecting auditory sources of interest.

STRF estimation. For frequency domain f, the input-output relation between a representation $S(f,t)$ of auditory input and the evoked cortical response $r(t)$ is modeled by a spectro-temporal response function (STRF) formulated as:

$$r_{\text{pred}}(t) = \sum_f \sum_\tau \text{STRF}(f,\tau) S(f, t-\tau) + \varepsilon(t)$$

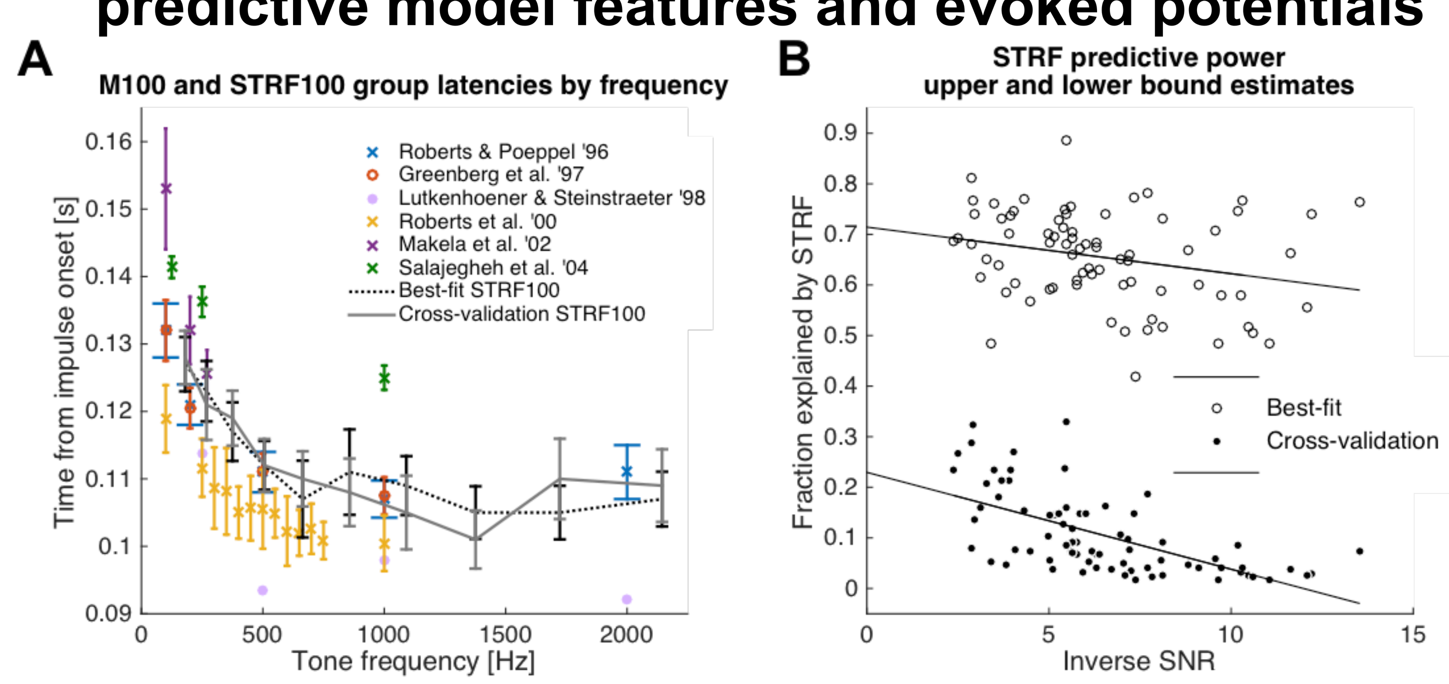
where $\varepsilon(t)$ is the residual of the evoked response not explained by the linear system.

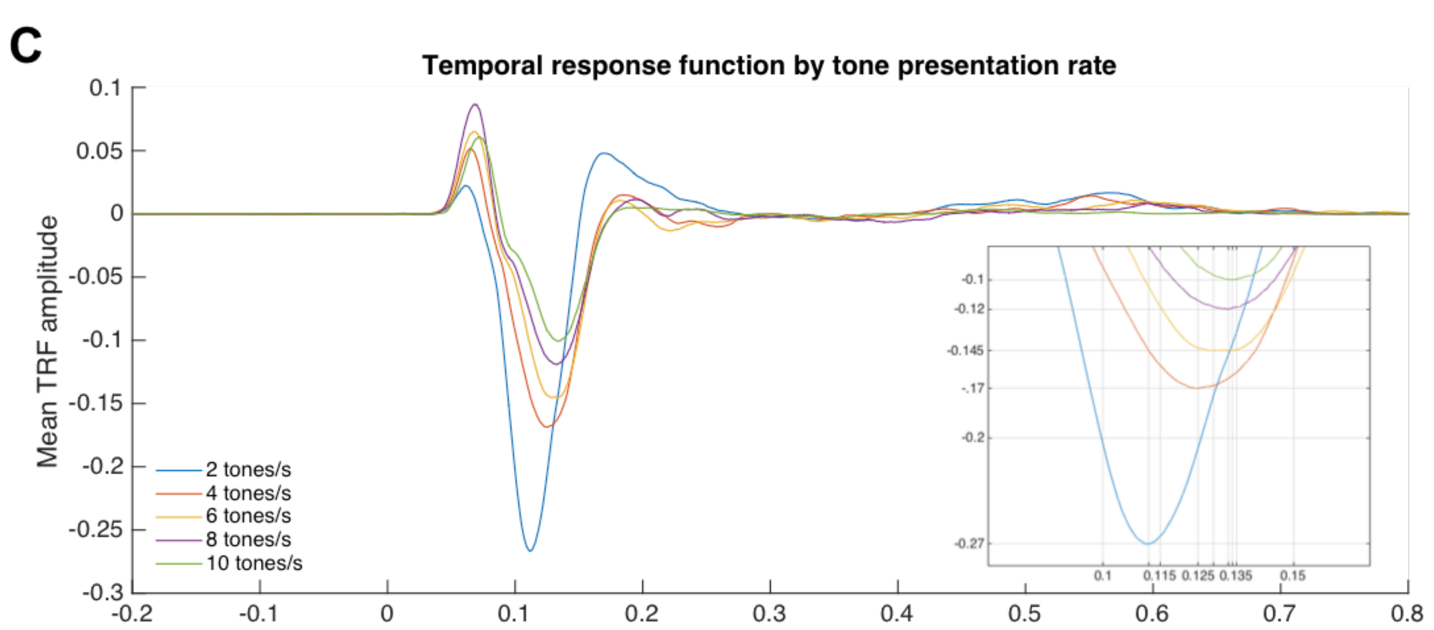
1 Spectrotemporal encoding models of MEG signals from human auditory cortex



(**A**) Group MEG spectro-temporal response functions following presentation of multitone random clouds include a positive-negative-positive complex, with tone cloud density changes introducing qualitative modifications to relative amplitude and delay. Average predictive ability across linear models comprises r correlation coefficients ranging between 0.31 and 0.38. (**B**) Group MEG response to the sparsest cloud shows deflections predicted by single tone onset response model (top). Empirical versus model responses based on STRFs optimized on self- or cross-validation (bottom).

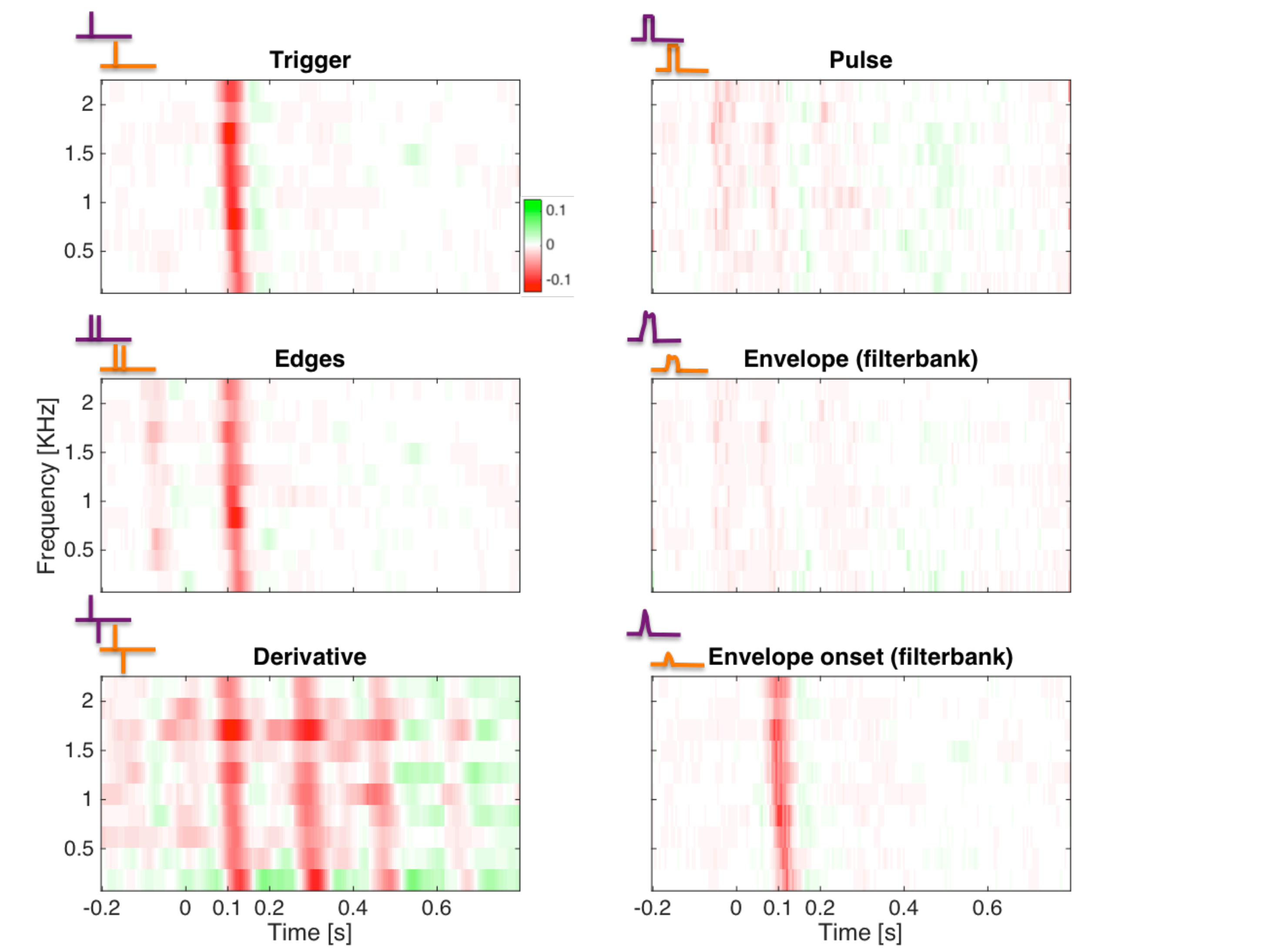
2 Consistency between response function predictive model features and evoked potentials





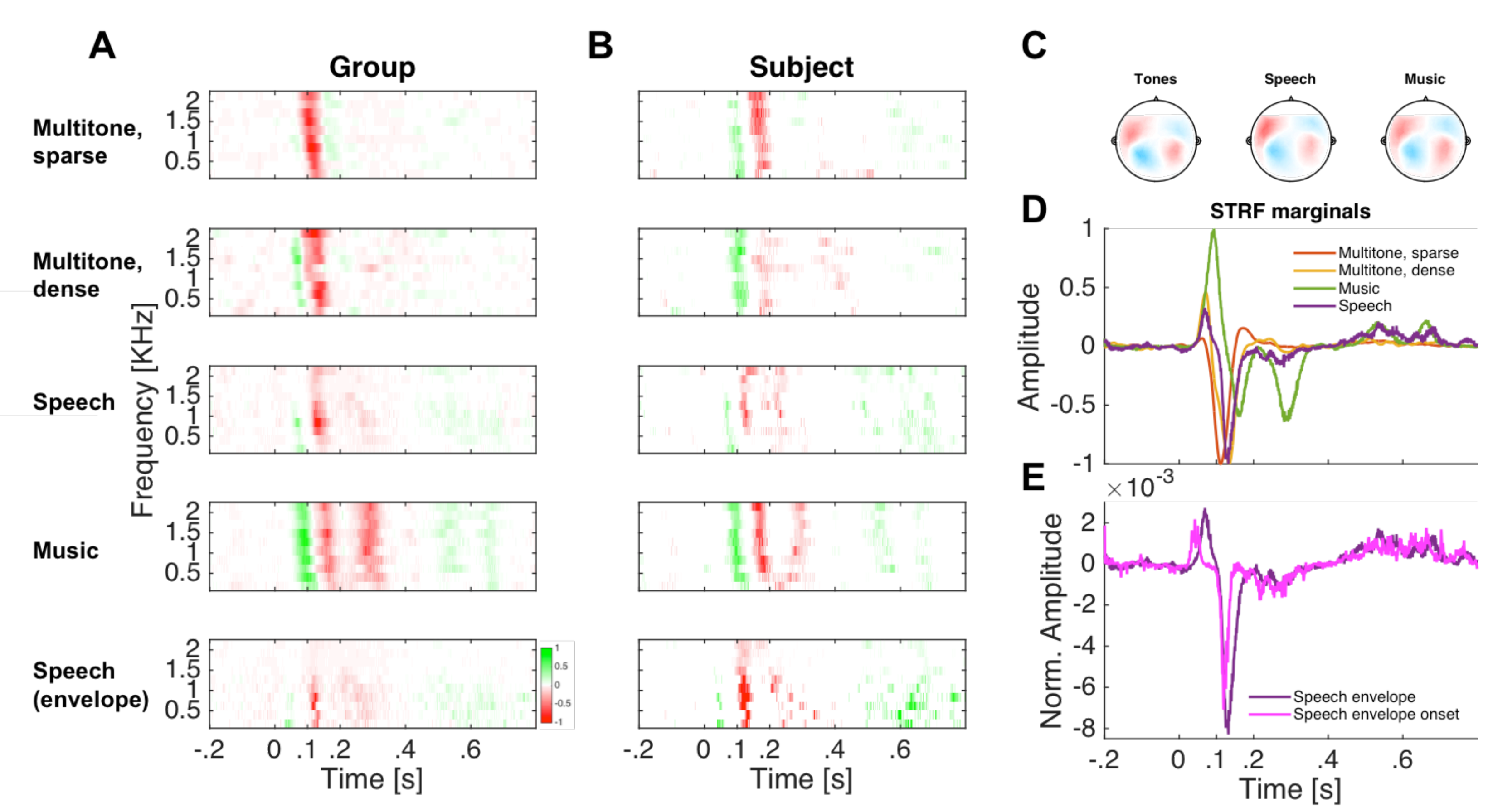
(**A**) Late-latency (~100 ms) negative peaks from the tone cloud MEG STRF become delayed by about 20 ms, as tone carrier frequency decreases from 2 to 0.2 KHz as with single tone evoked potentials. (**B**) Linear STRF models from MEG signal power may account for between 23% and 71% of explanatory power, by extrapolation the empirical model performance limit at the theoretical noise-less limit. (**C**) Temporal response functions are consistent with P1/P1m-N1/N1m-P2/P2m complex in evoked potentials.

3 Interpretational power across stimulus representations



Interpretational power revealed by group STRFs vary as stimulus is represented either by individual tone trigger times, timing and/or directionality of its temporal edges, or alternatively by tone pulse durations. Some of these abstract representations are consistent with standard representations of the acoustic envelope (and envelope onset) following a filterbank model.

4 Onsets as integrative features for STRF components across stimulus classes



(**A**) Group normalized STRFs across stimuli classes reveal considerable structure similarity when onset is extracted as a driving feature of the neuromagnetic response, in contrasts with datasets from MEG studies on speech (N=12) and music (N=15) processing (note groups different). (**B**) Neuromagnetic STRFs from same single participant in tones, speech, and music studies, again show remarkable structure consistencies when stimuli are represented by their temporal envelope onsets per frequency band. Component timing may differ according to both stimulus feature and class being modeled. (**C**) Data-driven virtual sensors from same subject as in (B) per study reveal strong scalp bihemispheric consistencies, with increased left hemisphere-bias for speech. (**D**) Across participants, timing of major neuromagnetic activity peaks, as shown by temporal marginals from (A), varies depending on stimulus class and/or context: earliest positive and negative deflections change with increasing acoustic density but also with additional spectrotemporal complexity as found in natural speech and music. (**E**) Group STRF marginals comparing both speech envelope and envelope onset activity. Timing differences are explainable by differential acoustic representation in early but not late (> 0.1s) activity peaks, suggesting attainment of higher order neural representation of elements in speech acoustics by 100 ms.

Conclusions

The spectro-temporal response function reveals stimulus features that are encoded in the neuromagnetic response, which it may predict.

STRF predictive power is commensurate with that from single/ multiunit electrophysiology.

Method may serve as extension of classic event-related potentials, gaining insight into aspects of responses that generalize or not across instance repetitions.

Method helps explore parsimonious representations of stimuli, and optimal for the interpretability of STRF kernel components.

Temporal resolution of MEG ideal to assess multi-stage auditory cortical processing of artificial and natural sound.

Acknowledgments

This work was funded by the National Institutes of Health (grant R01008342). We thank support to FCC from the Mexican Consejo Nacional de Ciencia y Tecnología.

CONACYT Consejo Nacional de Ciencia y Tecnología

NIH National Institute on Deafness and Other Communication Disorders (NIDCD)

References

David SV, Mesgarani N, Shamma SA (2007) **Estimating sparse spectro-temporal receptive fields with natural stimuli**. *Network*, 18(3): 191–212

de Cheveigné A, Simon JZ (2008) **Denoising based on spatial filtering**. *J Neurosci Methods*, 171(1):331–339.

Ding N, Simon JZ (2012) **Neural coding of continuous speech in auditory cortex during monaural and dichotic listening**. *J Neurophysiol*, 107(1):78-89.

Picton TW (2010) ***Human Auditory Evoked Potentials***. San Diego: Plural Publishing, Inc.

Sahani M, Linden JF (2003) **How linear are auditory cortical responses?** in *Advances in Neural Information Processing 15*, S Becker, S Thrun, K Obermayer, Eds., MIT Press, pp. 125–132.