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

The Computational Sensorimotor Systems Laboratory

NACS PROGRAM NEUROSCIENCE AND COGNITIVE SCIENCE

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.

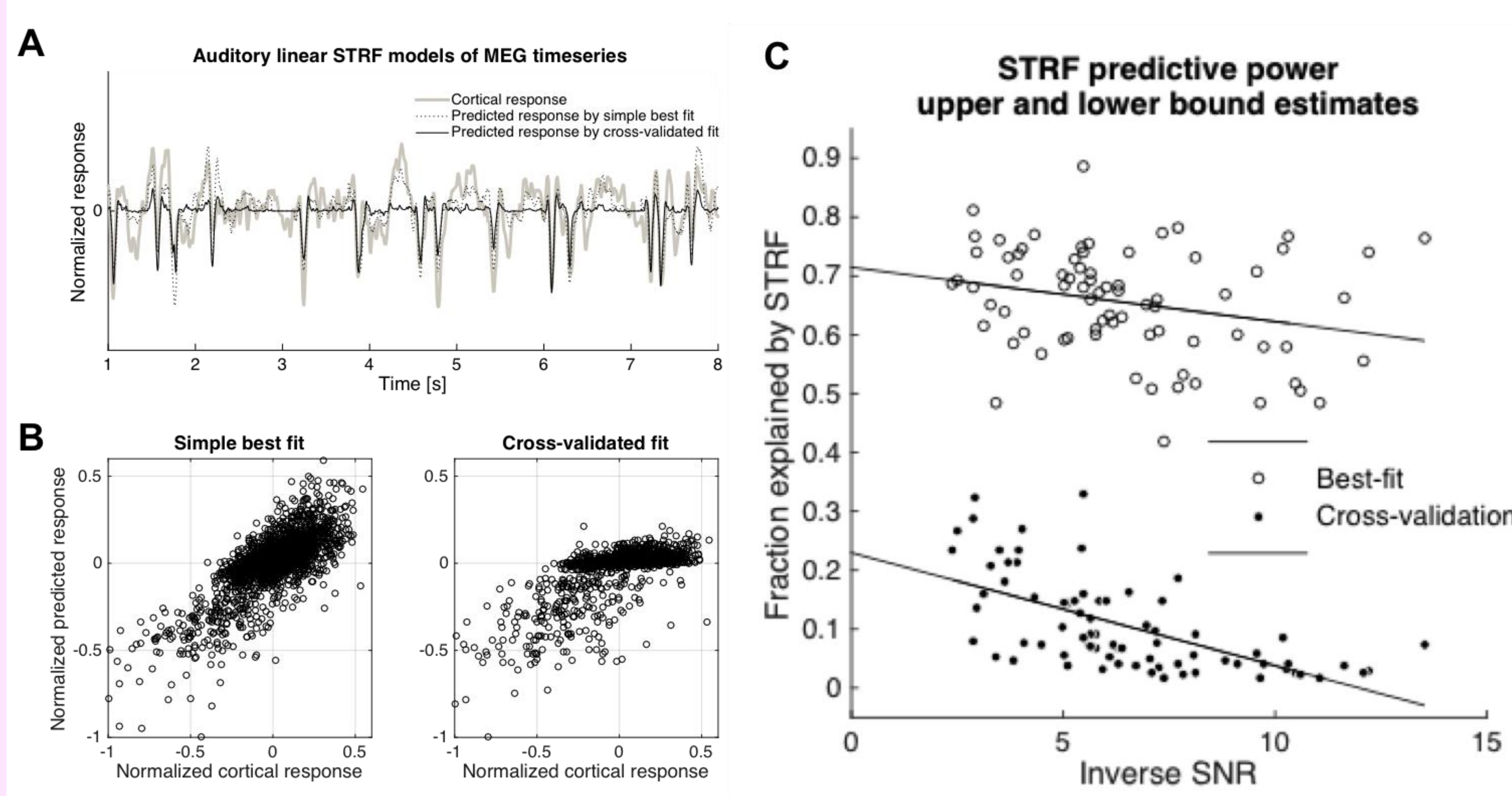
Data acquisition and analysis. Environmental and sensor noise contributions to neural signals from a 157-channel, whole-head MEG-KIT system (1 KHz sampling rate and 60 Hz notch filter) were estimated and removed. Sensor recordings were band-pass filtered 1-15 Hz then spatially filtered into a single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

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

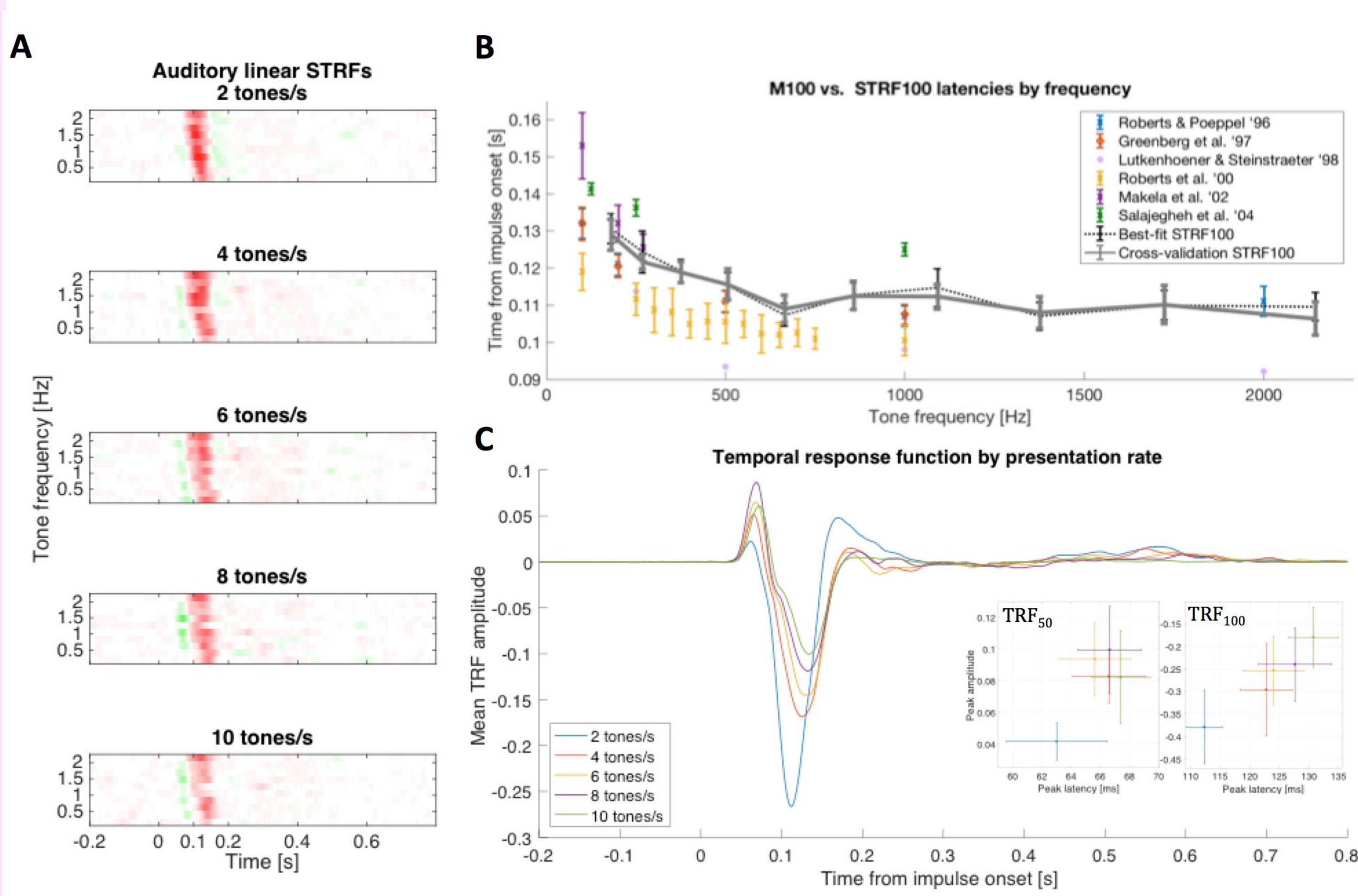
where $\varepsilon(t)$ is the residual of the evoked response not explained by the linear system.

1 Performance bounds to spectrotemporal encoding models of MEG signals from human auditory cortex



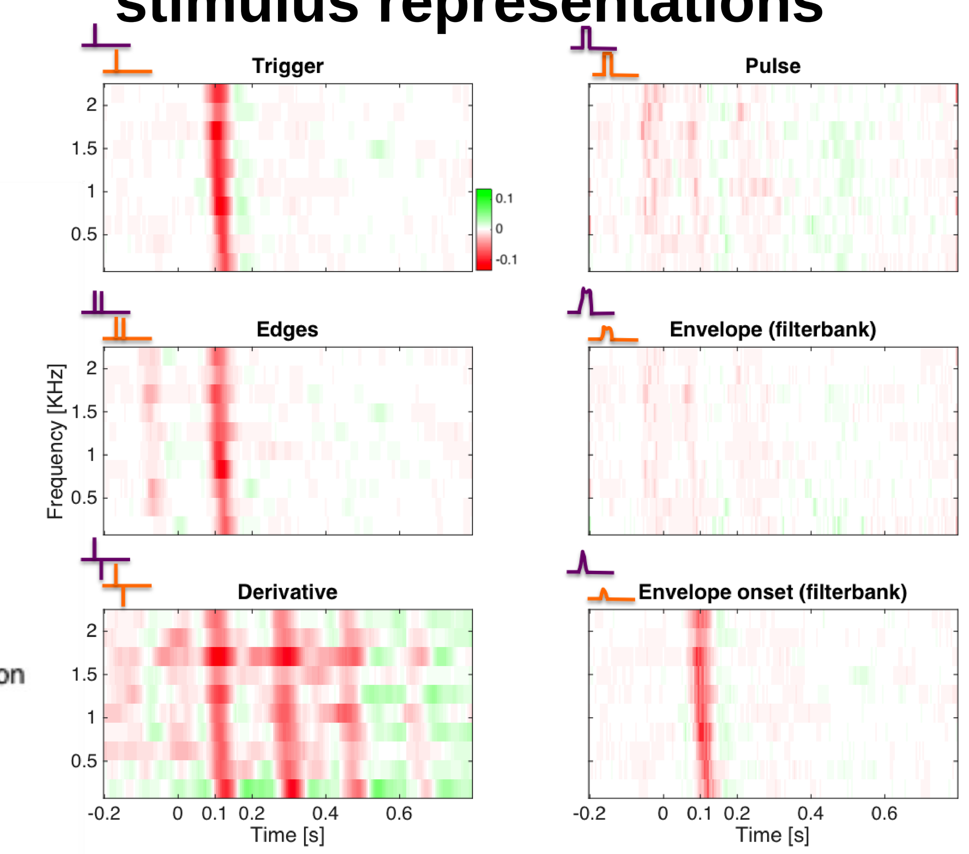
(**A**) Group MEG response following presentation of sparse multitone random clouds contains deflections predictable by linear STRF model. (**B**) Empirical versus model responses based on STRFs optimized on simple best fit or on cross-validation. (**C**) Linear STRF models from MEG signal power may account for between 23% and 71% of explanatory power, by extrapolation the empirical model performance at the noiseless theoretical limit.

2 Consistency between response function model features and evoked potentials



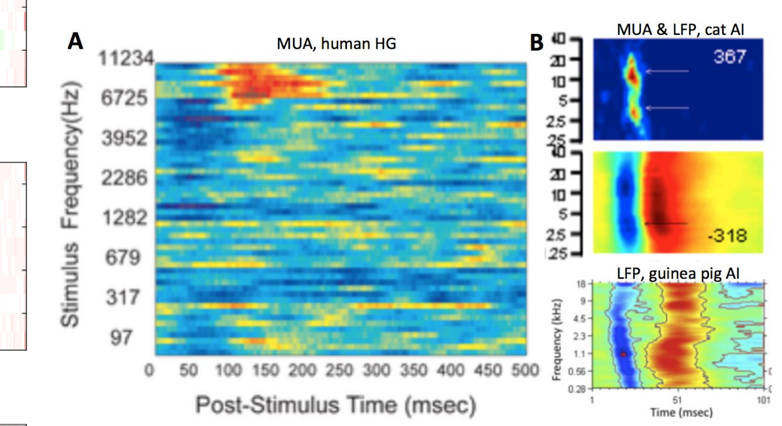
(**A**) Group MEG STRFs following presentation of multitone random clouds of different density feature a positive-negative-positive complex, with relative amplitude and delay dependencies on tone cloud density. Correlation coefficients range: r 0.31 to 0.38. (**B**) Late-latency (~100 ms) negative peaks from the tone cloud MEG STRF delayed by about 20 ms, as tone carrier frequency decreases from 2 to 0.2 KHz as with single tone evoked potentials, as in classic evoked potential studies. (**C**) Temporal response functions are consistent with P1/P1m-N1/N1m-P2/P2m complex in evoked potentials, with modulations to individual components' amplitude and latency.

3 Interpretational power across stimulus representations



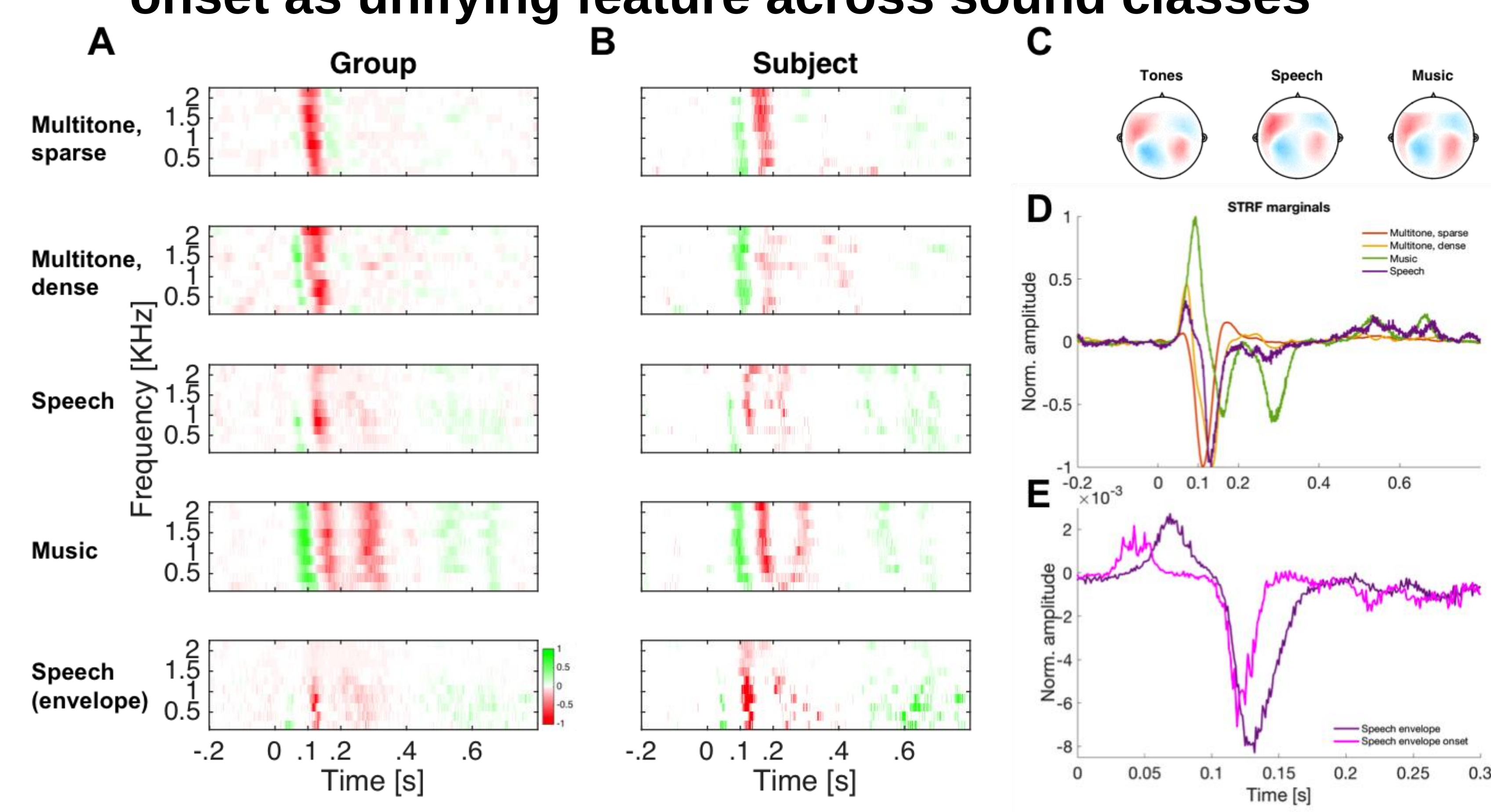
Interpretational power of group STRFs leveraged by stimulus representation: individual tone trigger times, timing and/or directionality of temporal edges, or tone pulse duration. Some of abstract representations consistent with standard acoustic envelope (and envelope onset) following a real filterbank model applicable to natural sounds.

5 Comparisons to electrophysiology-derived STRFs



Gammatone cloud STRF reproductions from single electrode recordings in primary auditory cortex. (**A**) Single unit activity from awake human (20 tones/s; Jenison et al., 2015). (**B**) Multiunit activity (top) and local field potential (centre) from anaesthesized cat (4 tones/s; Noreña et al., 2008); LFP from guinea pig (bottom; ~2 tones/s; Gaucher et al., 2011).

4 Insights into differential processing of acoustic onset as unifying feature across sound classes



(**A**) Group-normalized STRFs across stimuli classes reveal considerable structure similarity when onset is extracted from tone cloud stimulus, but also speech (N=12) and music (N=15) stimuli (separate subject pools).
(**B**) Subject STRFs from same listener in tones, speech, and music studies, again show remarkable structure consistencies when stimuli are represented by their temporal envelope onsets per frequency band. Latencies in elements in STRF timing may differ per stimulus feature and class.
(**C**) Data-driven MEG virtual sensor maps from same subject as in (B) per study reveal strong scalp bihemispheric consistency.
(**D**) Time marginals from (A): components' latency by stimulus class and/or context. Amplitudes are normalized.
(**E**) Time marginals from (A): : components' latency by stimulus feature - envelope vs. envelope onset. Timing differences explainable by differential acoustic representation in early (~50ms) but not late (> 0.1s) activity peaks, consistent with 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. Thanks to Elizabeth Camenga, Katya Dombrowski and Benjamin Walsh for the music data, and to Marisel Villafañe-Delgado for the speech data.

CONACYT Consejo Nacional de Ciencia y Tecnología

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

References

David SV, et al. (2007) *Network*, 18(3): 191–212

de Cheveigné A, Simon JZ (2008) *J Neurosci Methods*, 171(1):331–339

Ding N, Simon JZ (2012) *J Neurophysiol*, 107(1):78–89

Gaucher Q, et al. (2012) *J Physiol-Paris*, 106:93–103

Jenison RL, et al. (2015) *PloS One*, 10:e0137915

Noreña AJ, et al. (2008) *J Neurosci*, 28:8885–8896

Sahani M, Linden JF (2003) *Advances in Neural Information Processing 15*, MIT Press, pp. 125–132.