Neural source dynamics of brain responses to continuous speech: from acoustics to comprehension

Christian Brodbeck*[1], Alessandro Presacco[2] & Jonathan Z. Simon[1,3,4] *brodbeck@umd.edu

[1]Institute for Systems Research, [3]Department of Electrical and Computer Engineering, [4]Department of Biology, University of Maryland, College Park, Maryland; [2]Department of Otolaryngology, University of California, Irvine, California

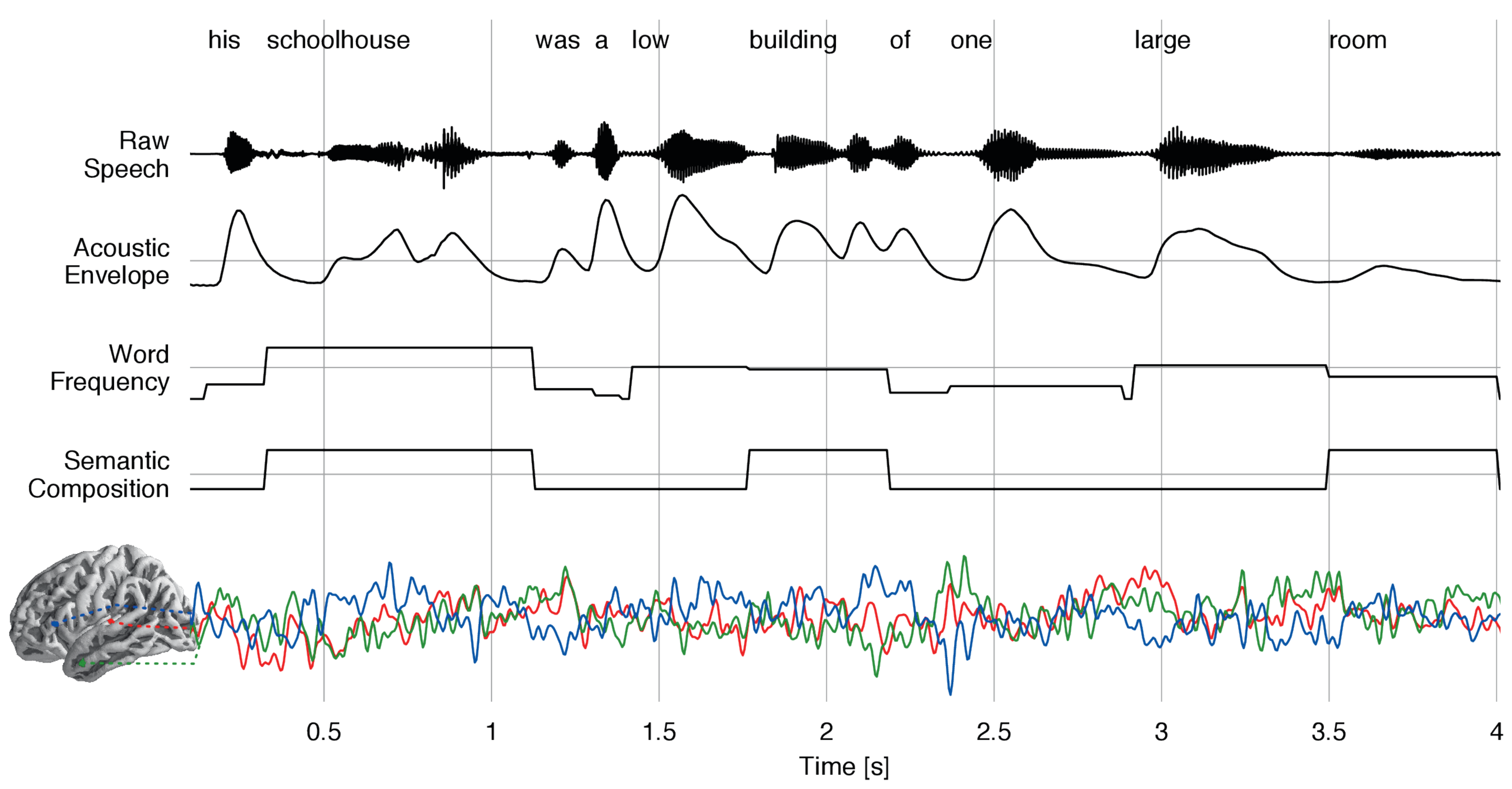
Computational Sensorimotor Systems Lab

Introduction

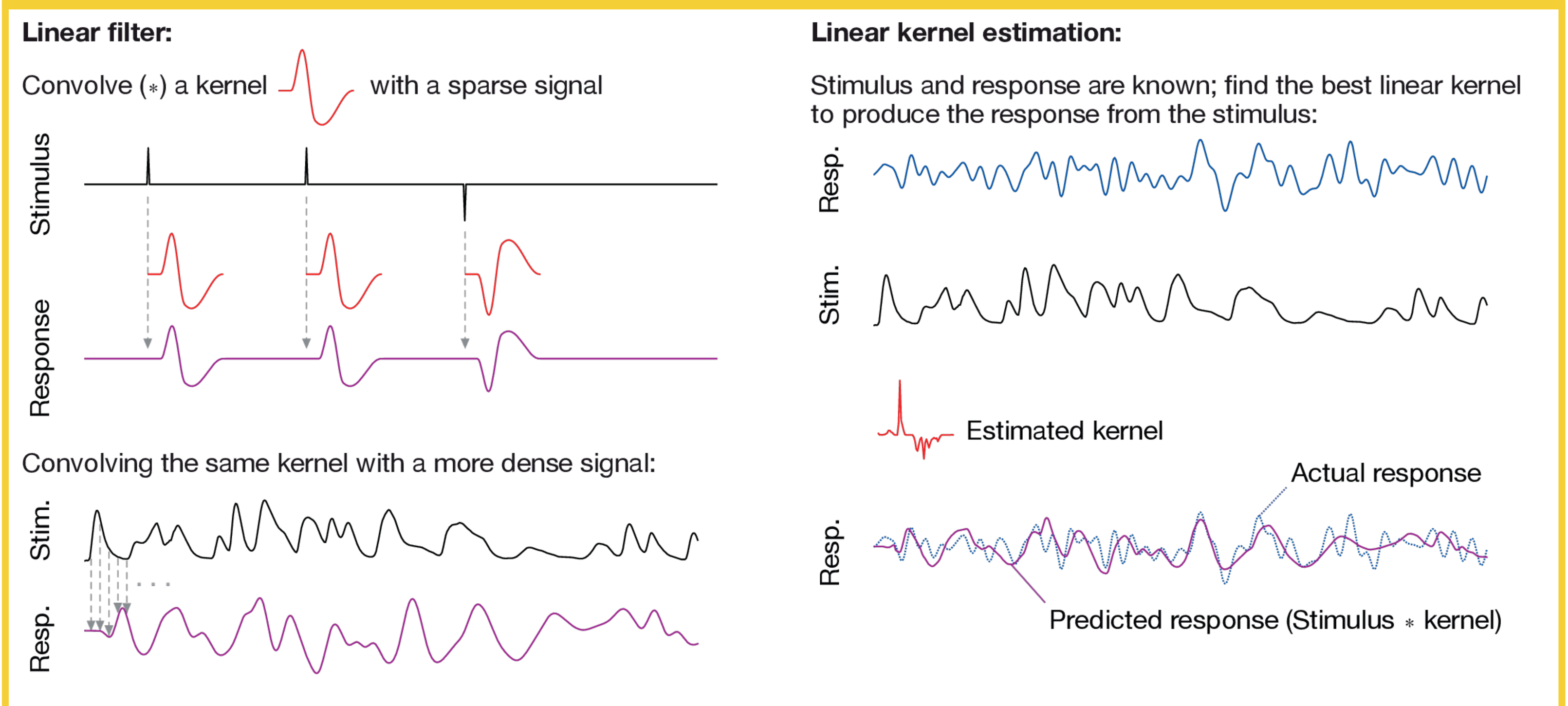
The high temporal resolution of electro- and magnetoencephalography (EEG/MEG) makes them ideal tools to study brain responses to rapidly evolving continuous stimuli such as speech. Linear kernel estimation has been used to deconvolve EEG and MEG responses to continuous stimuli (see box "Linear kernel estimation"). However, this analysis is typically applied to sensor space data, not using the full neural source localization power of MEG. To localize responses anatomically, we computed distributed minimum norm source current estimates of continuous MEG data and estimated a separate response function for each virtual current source dipole.

We analyzed data from participants listening to excerpts from an audiobook with 3 predictor variables:

- **Acoustic envelope**: acoustic power across frequency bands
- **Word frequency**: strong predictor of lexical processing
- **Semantic composition**: estimate of semantic integration; correlated with other comprehension-related variables



Method: linear kernel estimation



Methods

MEG data

- 17 participants listened to 2 one-minute long segments from a narration of The Legend of Sleepy Hollow by Washington Irving; each segment was repeated 3 times for a total of 6 minutes
- An average brain model ("fsaverage", FreeSurfer) was scaled and coregistered to each subject's head shape
- MEG data were projected to source space using distributed minimum norm inverse solution (approximately 5000 virtual source dipoles, regularly spaced on the white matter surface, oriented perpendicular to the cortical surface)

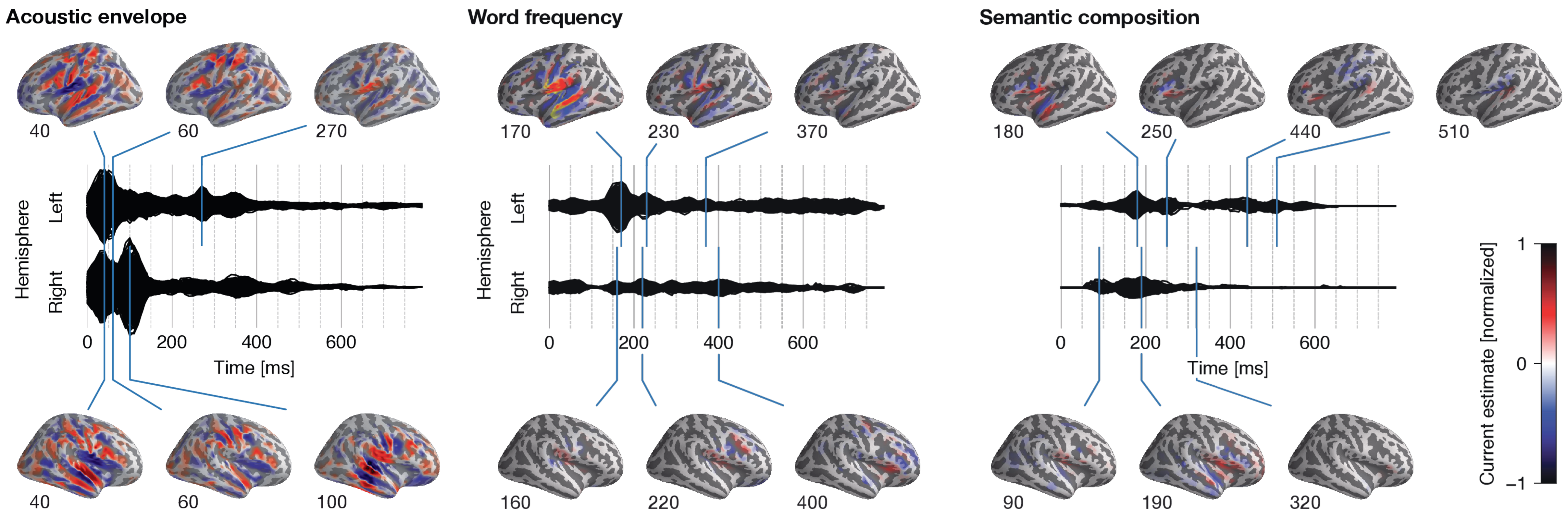
Predictor variables

- Acoustic envelope: average of all frequency channels of an auditory brainstem model (Yang et al., 1992)
- Word frequency: log frequency values from the SUBTLEX database (Brysbaert and New, 2009)
- Semantic composition: Content words matching any of the patterns of semantic composition analyzed by Westerlund et al. (2015) were marked as 1, all other time points as 0

Response functions

- Response functions were estimated separately for each virtual current source dipole using the boosting algorithm (David et al., 2007)
- The response functions were assessed for spatio-temporal patterns that differed significantly from zero using spatio-temporal permutation tests based on threshold-free cluster enhancement (Smith and Nichols, 2009)

Results



- Auditory cortex (~40 ms, ~100 ms)
- Sensorimotor parietal and frontal cortices (~50 ms)

- Strong left-lateralized response in auditory cortex (~170 ms)
- Later, weaker bilateral frontal response

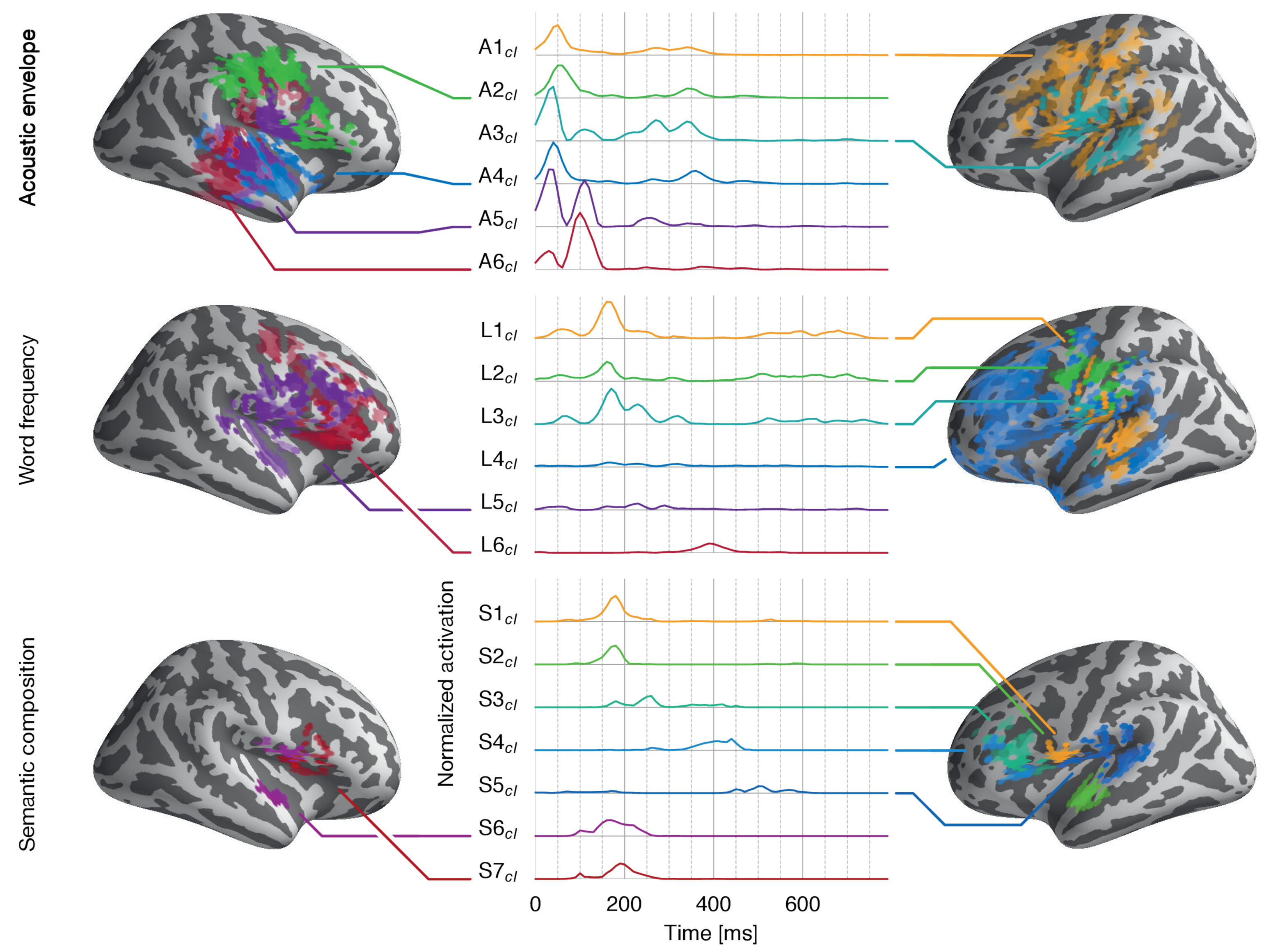
- Left hemisphere: temporal progression from anterior temporal lobe to inferior frontal gyrus activation
- Right-hemisphere: similarly localized, temporally more diffuse

Response functions

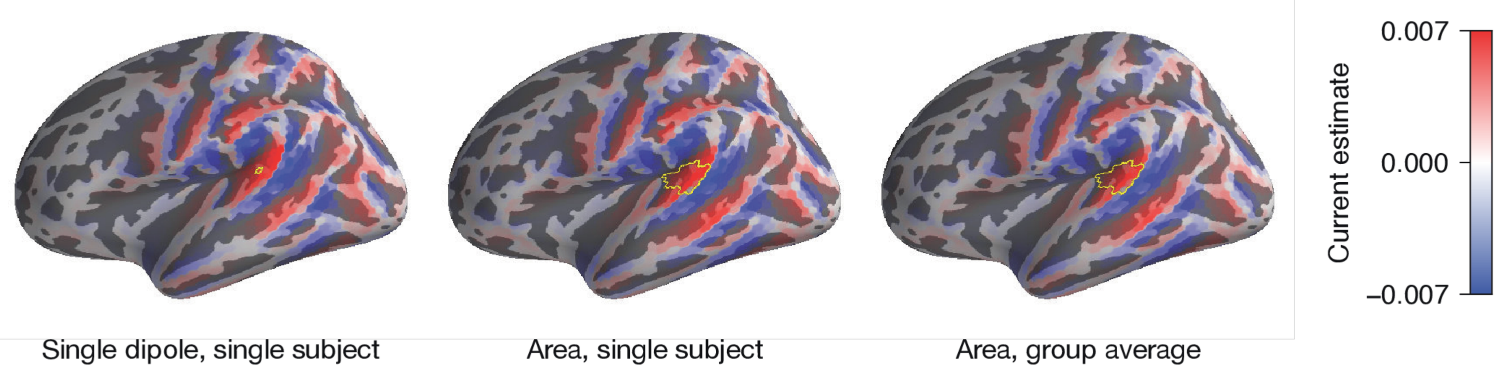
Average response functions across subject ($p < .05$). Non-significant values were set to zero.

Clustered response functions

Because of the smoothness of MEG source estimates (see box "Point spread function") response functions are composed of multiple overlapping responses. However, responses due to the same underlying neural source should exhibit the same time course. Hierarchical clustering (Ward, 1963) of dipoles based on their response time-course revealed separable neural sources for each predictor variable.



Point spread function



The theoretical accuracy of MEG source localization can be evaluated using the point spread function. Since both the forward model *L* and the inverse operator *G* are linear matrix operations, the source estimate of a hypothetical source current vector *j* can be computed by combining both operations, $G \cdot L \cdot j$. The source estimate for a hypothetical point source is the "point spread function". Hypothetical sources are indicated by the yellow outline.

Conclusions

- Combining linear kernel estimation with source localization allows anatomically separating brain responses to different stimulus properties
- Localization preserves temporally precise response functions (order of tens of milliseconds)
- Simultaneously sensitive to variables related to higher cognitive levels in speech comprehension as well as basic acoustic properties
- Robust responses from just 6 minutes of data
- Broadens the possibilities for studying speech comprehension with natural stimuli
- Applicable also to other continuous stimuli

References

Brysbaert, M., and New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behav Res Methods 41, 977–990.

David, S.V., Mesgarani, N., and Shamma, S.A. (2007). Estimating sparse spectro-temporal receptive fields with natural stimuli. Netw. Comput. Neural Syst. 18, 191–212.

Smith, S.M., and Nichols, T.E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage 44, 83–98.

Ward, J.H. (1963). Hierarchical Grouping to Optimize an Objective Function. J. Am. Stat. Assoc. 58, 236–244.

Westerlund, M., Kastner, I., Al Kaabi, M., and Pylkkanen, L. (2015). The LATL as locus of composition: MEG evidence from English and Arabic. Brain Lang 141, 124–134.

Yang, X., Wang, K., and Shamma, S.A. (1992). Auditory representations of acoustic signals. IEEE Trans. Inf. Theory 38, 824–839.

Supported by National Institutes of Health (NIH) grant R01-DC-014085 and ICAC

Poster PDF available at http://ter.ps/simonpubs
Manuscript on bioRxiv: https://doi.org/10.1101/182881