Neural source dynamics of brain responses to continuous stimuli with MEG: speech processing from acoustics to comprehension

Christian Brodbeck*[1], Alessandro Presacco[2] & Jonathan Z. Simon[1,3,4,5]
[1]Institute for Systems Research, [3]Neuroscience and Cognitive Science Program, [4]Department of Electrical and Computer Engineering, [5]Department of Biology, University of Maryland, College Park, Maryland
[2]Department of Otolaryngology, University of California, Irvine, California
*brodbeck@umd.edu

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 then used permutation-based tests and hierarchical clustering to find significant spatio-temporal patterns in the responses. To demonstrate this method, we analyzed the MEG responses of participants listening to segments of a story. We used predictor variables representing different processing steps in comprehension to show differences in anatomical localization.

Methods

MEG Data

- 17 Participants listened to 2 one-minute long segments from a narration of *The Legend of Sleepy Hollow* by Washington Irving, read by a male speaker in quiet background at ~70 dB SPL; 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
- Raw MEG data were preprocessed using temporal signal space projection (Taulu and Simola, 2006) and band-pass filtered 1-40 Hz
- MEG data epochs relative to stimulus onset, downsampled to 100 Hz
- Epoched 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

- Phoneme boundaries were marked using the Gentle forced aligner and manually adjusted using Praat
- An acoustic envelope representation was computed as the average of all frequency channels of a model of the acoustic transformations performed by the auditory brainstem (Yang et al., 1992)
- Word frequency was coded using log frequency values from the SUBTLEX database (Brysbaert and New, 2009), with higher values reflecting less frequent words
- 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

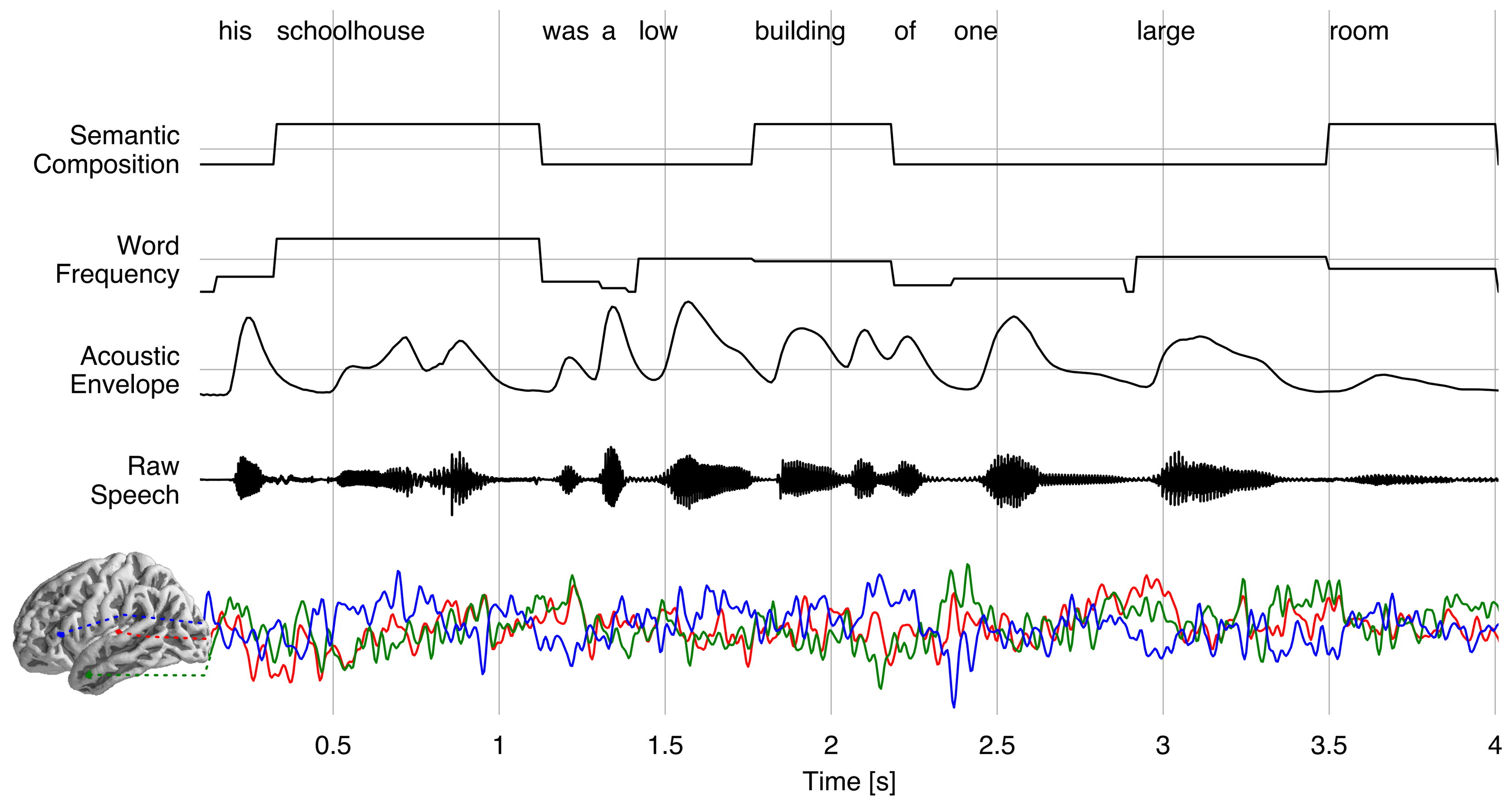
- Response functions were estimated separately for each virtual current source dipole using the boosting algorithm (David et al., 2007)
- The response functions were assessed for spatio-temporal patterns that differed significantly from zero using spatio-temporal permutation tests based on threshold-free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

Clustering of response functions

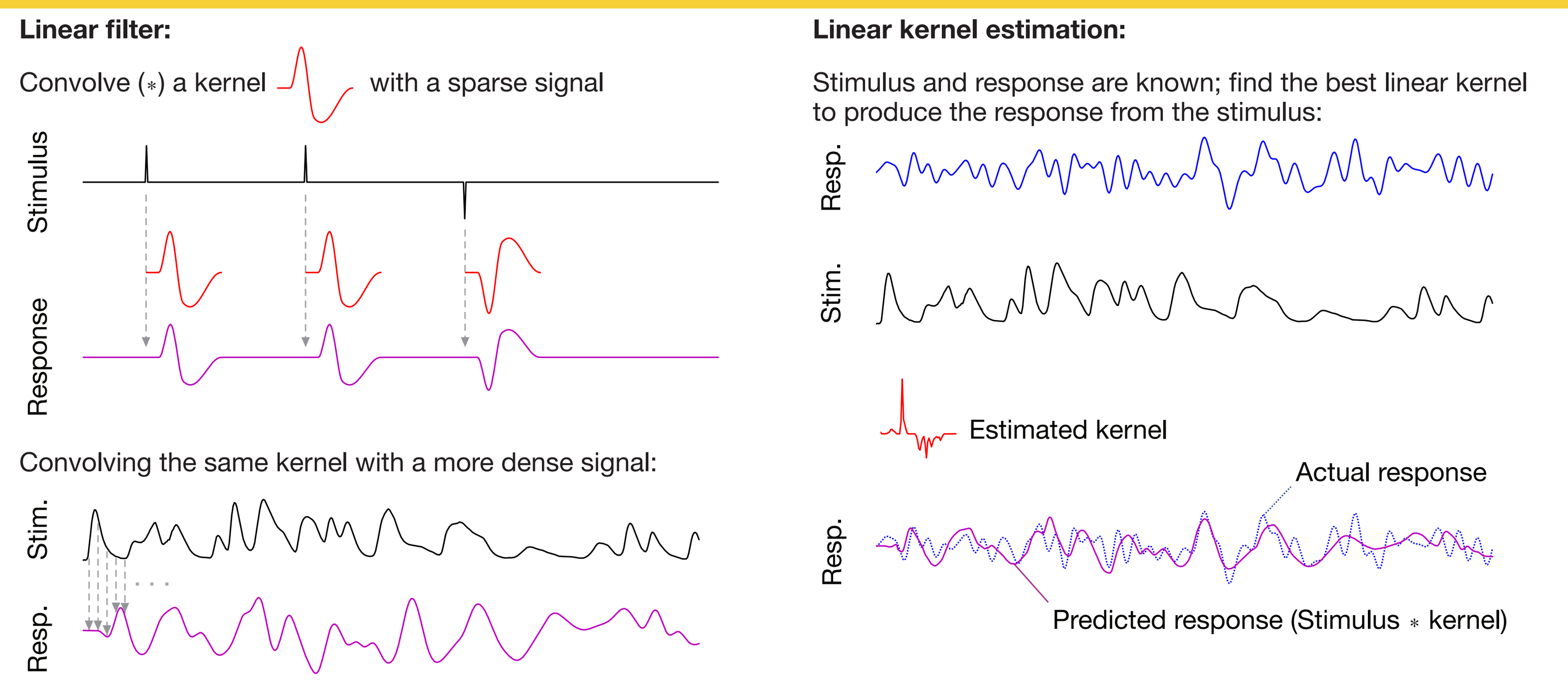
- Significant response functions were grouped using hierarchical clustering, minimizing the sum squared error (Ward, 1963)
- Starting with a single cluster, clusters were subdivided until the next split would reduce the error by less than 1% of the total sum squared
- Clusters that combined low amplitude sources of "halos" of one or more stronger clusters were visually identified and removed

Analysis

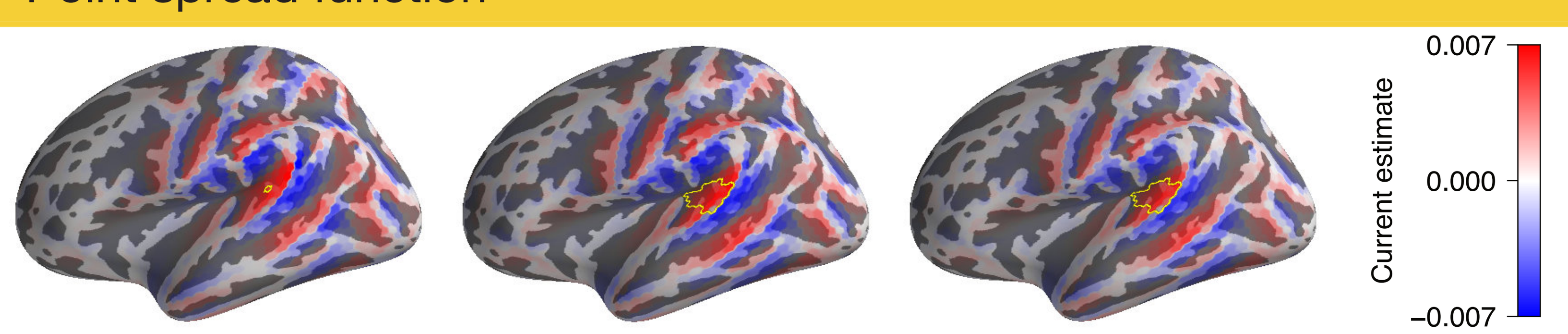
Predictor variables:
- **Acoustic envelope**: acoustic power across frequency bands
- **Word frequency**: less frequent words associated with larger values
- **Semantic composition**: estimate of the amount of semantic integration, but correlated with other comprehension-related variables



Linear kernel estimation



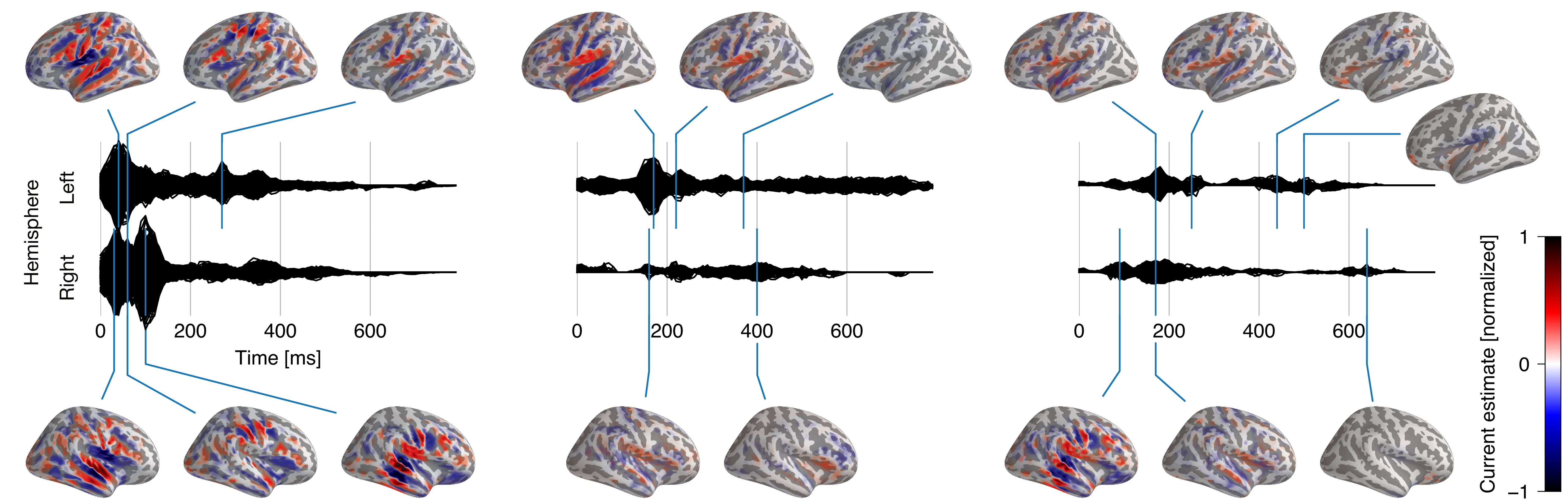
Point spread function



Single dipole, single subject — Area, single subject — Area, group average

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.

Results



Acoustic envelope
- Bilateral early response
  - auditory cortex (~30 ms)
  - sensorimotor parietal and frontal cortices (~50 ms)
- Right-lateralized later response
  - auditory cortex (~100 ms)

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

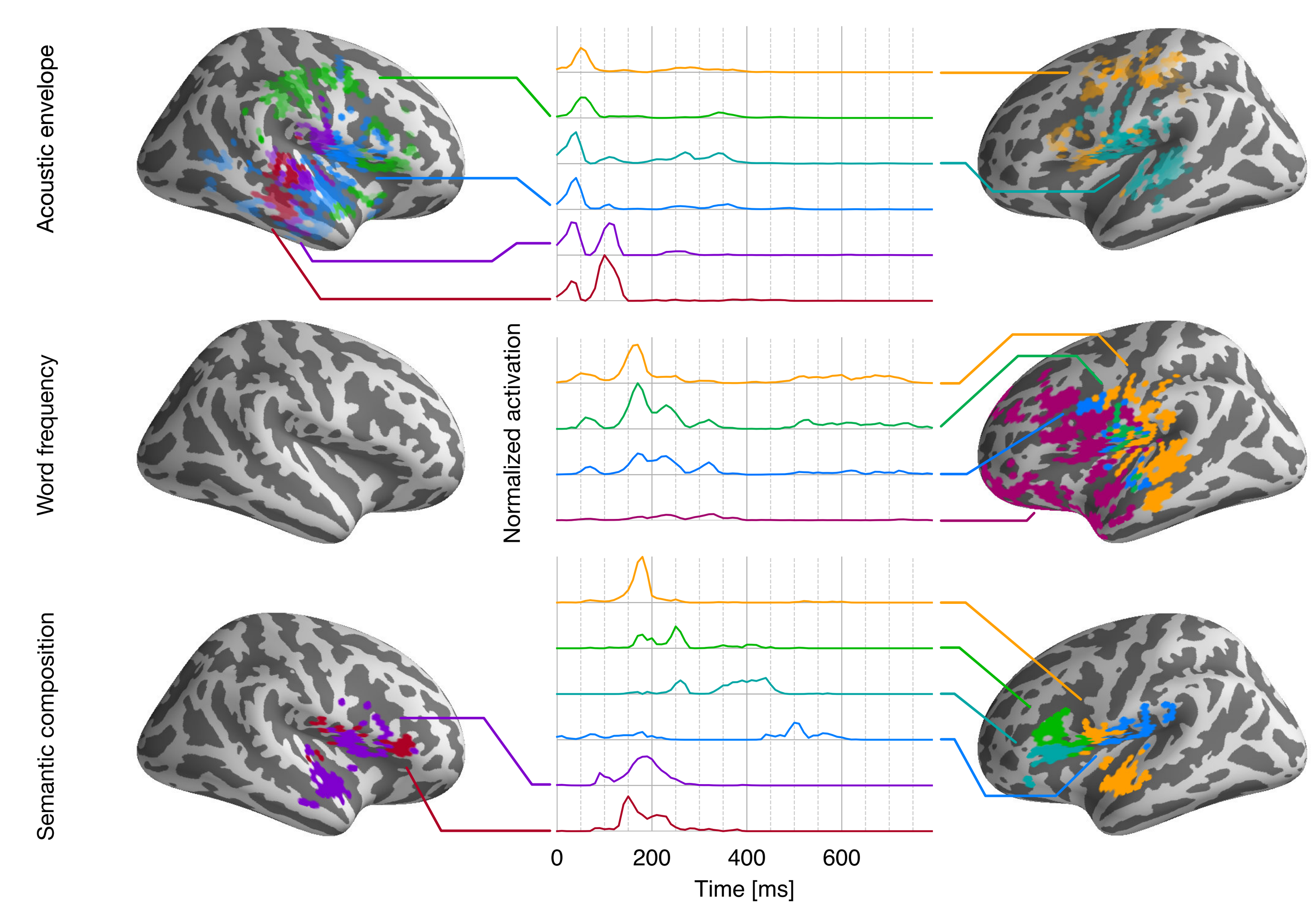
Semantic composition
- Left hemisphere:
  - temporal progression from anterior temporal lobe to inferior frontal gyrus activation
- Right-hemisphere
  - similarly localized, temporally more diffuse

Response functions

Average response functions across subject. Each black line reflects a single virtual current dipole. Non-significant values were set to zero for visualization.

Clustered response functions

Because of the smoothness of MEG source estimates (see box "Point spread function") response functions are composed of multiple overlapping responses. To find independent sources we used hierarchical clustering of dipoles based on their time-course (separately for each predictor variable).



Discussion

- Results confirm viability of analyzing continuous stimuli
- 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.

Taulu, S., and Simola, J. (2006). Spatiotemporal signal space separation method for rejecting nearby interference in MEG measurements. Phys. Med. Biol. 51, 1759.

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 SAND
PDF available at https://www.isr.umd.edu/Labs/CSSL/simonlab/Publications.html