Spatiotemporal dynamics of speech comprehension as revealed by MVPA

Caroline M. Whiting[1,2], Alexandra Woolgar[2,3], Li Su[1,2], Elisabeth Fonteneau[1,2], & William D. Marslen-Wilson[1,2]

[1]Department of Psychology, University of Cambridge; [2]MRC Cognition and Brain Sciences Unit, Cambridge; [3]ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, Australia

Introduction

Extensive evidence has revealed a bilateral fronto-temporal network supporting speech comprehension [1], and a left-lateralised sub-system specialised for processing of grammatical complexity (e.g. *walk-ed*; [2]).

Using combined MEG-EEG, we addressed how neural activity in these networks is modulated by grammatical properties of the speech input. We manipulated the presence of an inflectional affix (*–s*, *-ed*) to investigate what regions are involved in grammatical computations. Increased LH engagement was predicted for forms containing an inflectional ending.

We applied multivariate representational similarity analysis (RSA; [3, 4]) to examine the information carried by patterns of activity across multiple voxels, and to assess what linguistic dimensions are being processed in different regions of the neural language system.

Methods

Stimuli

Verb phrases were presented in three contexts, using 10 regular and 10 irregular verbs. Each item was repeated 12 times in order to compute a stable estimate for items-level analyses.

	regular	irregular
no affix	I walk	I fall
-s affix	He walks	He falls
past-tense	He walked	He fell

An acoustic baseline (Musical Rain) was created for each item, derived by jittering formants in each sound file. Musical Rain (MR) shares the acoustic properties of speech but is not interpretable.

Twenty right-handed, native English speakers took part in the study. They performed an occasional (10%) one-back semantic completion task, in which they were required to decide if the verb phrase was coherent within a sentence.

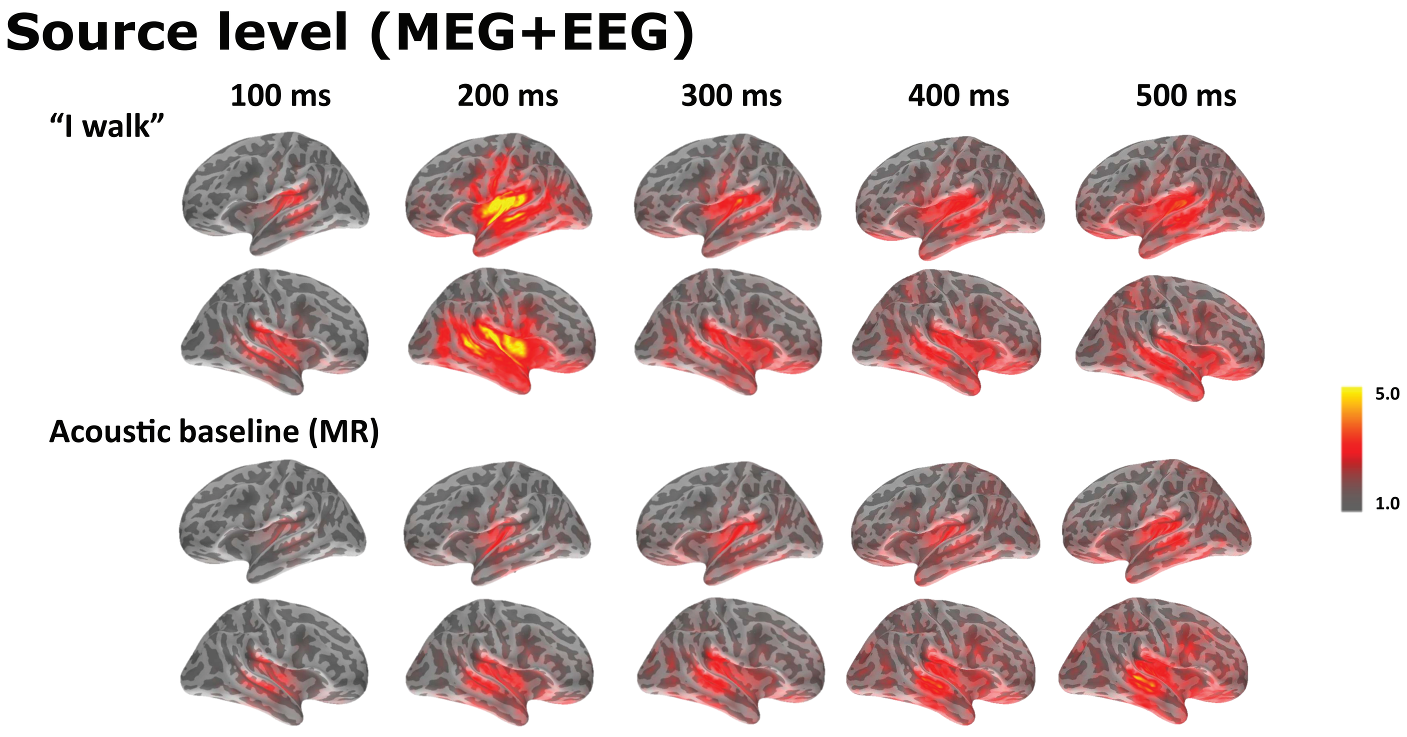
Acquisition and Source Reconstruction

Concurrent MEG-EEG data were acquired from a 306-channel Vectorview system with a 70-channel EEG cap. Epochs were generated from -100 to 700 ms from the onset of the sound file, and from -300 to 200 ms from the onset of the affix (-s/-ed).

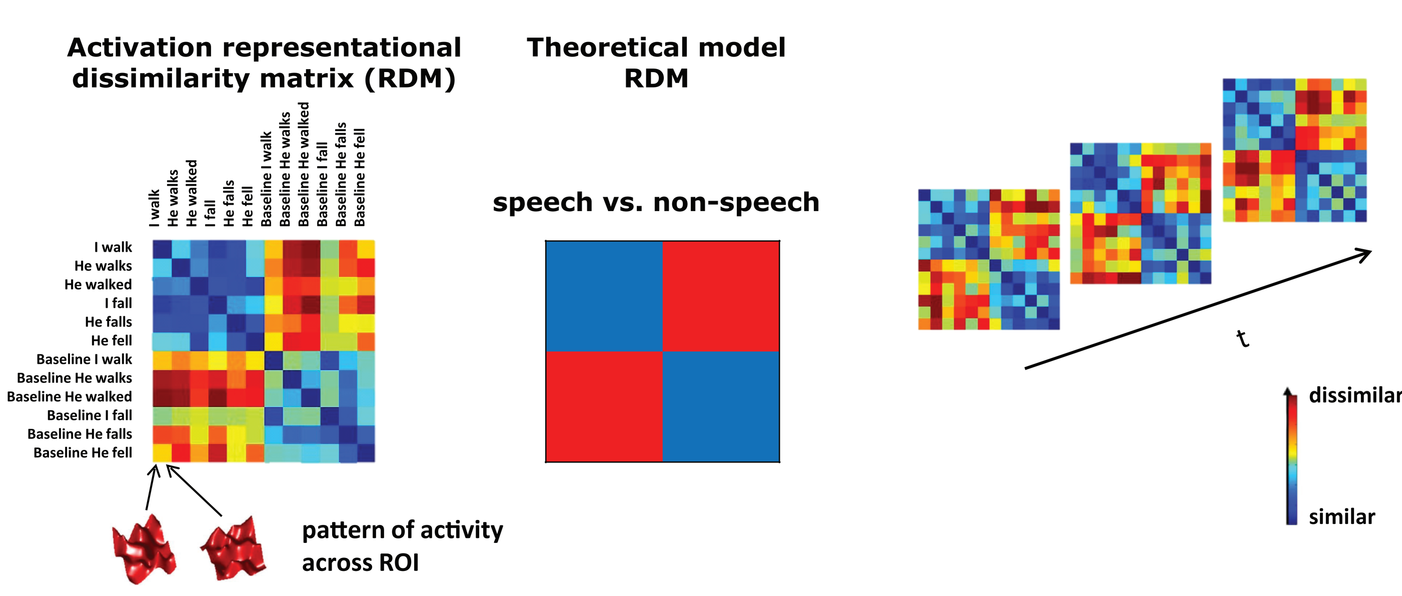
A three-layer boundary element model was created using FreeSurfer from individual structural MRIs. L2 minimum norm estimation (MNE) was used to compute MEG+EEG solutions. Regions of interest were defined anatomically in FreeSurfer.

Representational Similarity Analysis

Source level (MEG+EEG)



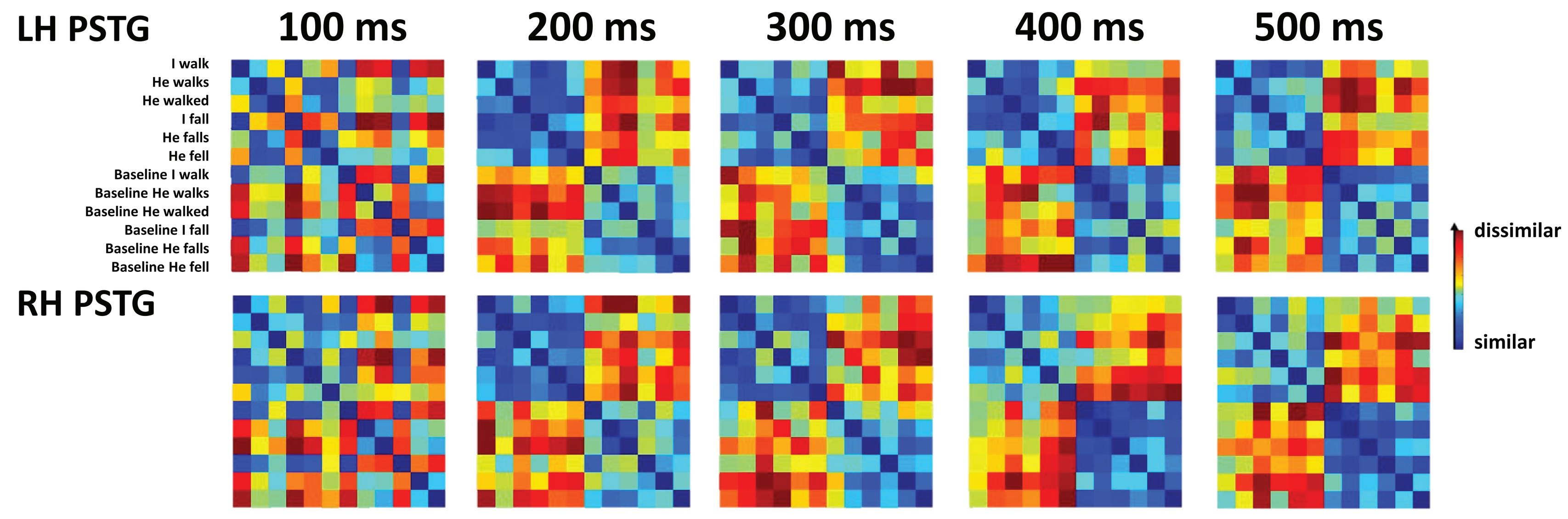
Mean source activity (noise-normalised MNE) for non-affixed condition (I walk) and matched Musical Rain from 100 to 500 ms. Increased activation appeared in speech-related areas for words compared to the Musical Rain baseline.



Patterns of source-level activity are expressed as representational dissimilarity matrices (RDMs), showing the correlation distance (1-correlation value) between pairs of conditions. Activation RDMs are compared to theoretical model RDMs, which hypothesise the stimulus distinctions in a region, allowing us to investigate specific language processing predictions. The added dimension of time provides one matrix at every time point.

MEG/EEG Results

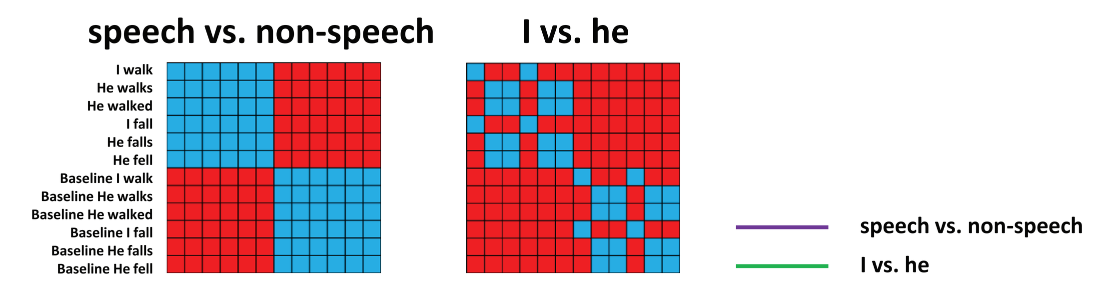
Data RDMs

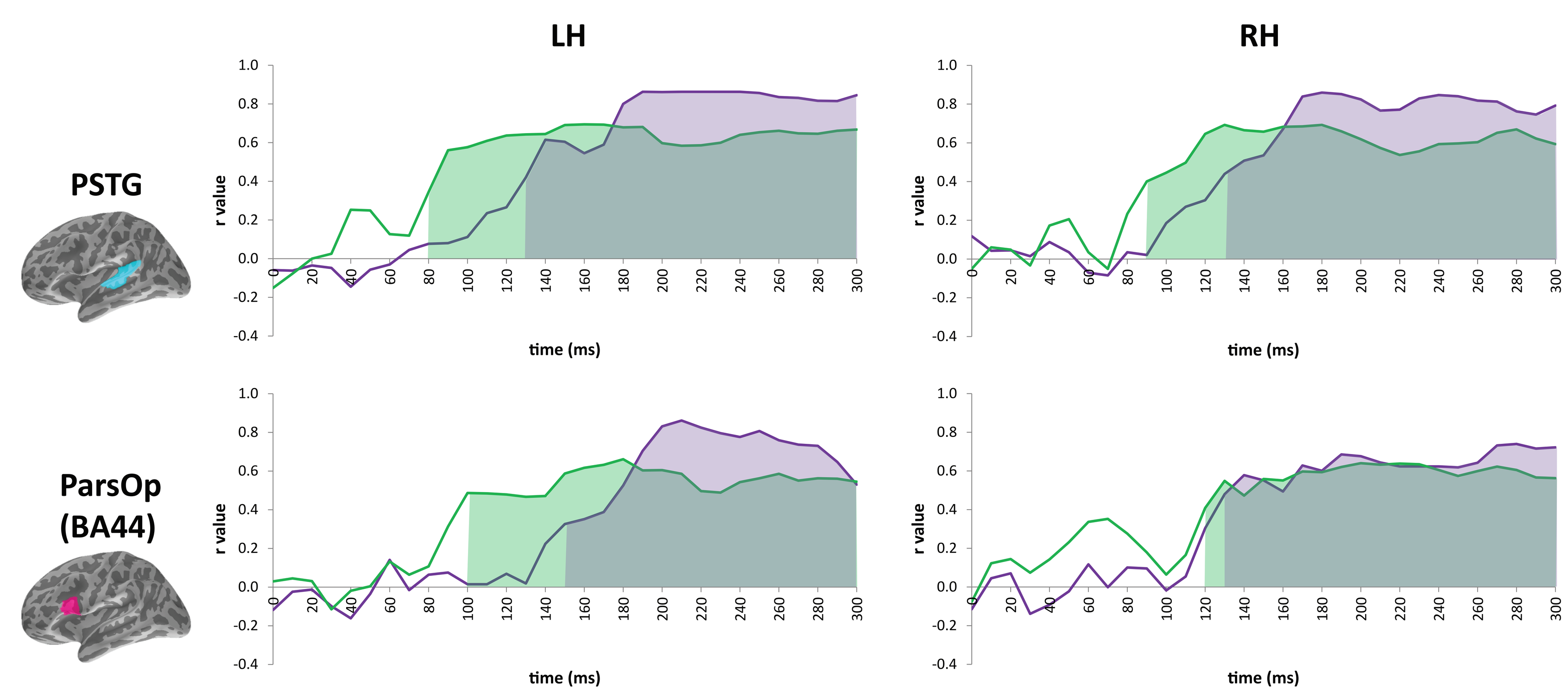


Mean activation RDMs for six word and six Musical Rain conditions in bilateral posterior superior temporal gyrus (PSTG), showing a strong dissociation between words and the Musical Rain baseline

Analyses aligned to speech onset

Twelve conditions (words and Musical Rain), averaged over 10 items in each condition. Analyses are aligned to the onset of the speech file. Shaded regions show a significant fit between data and model at p < .01.

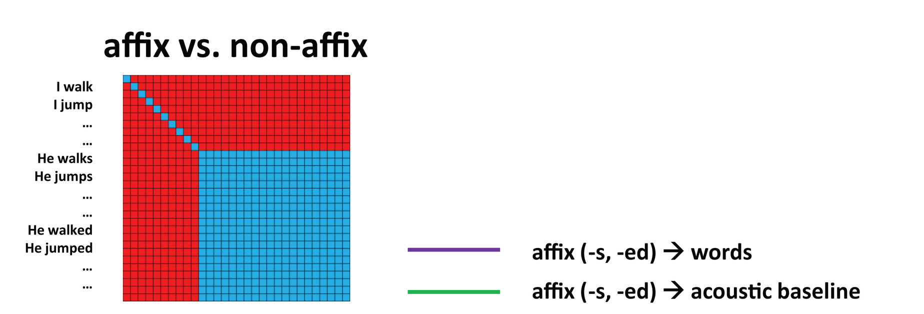


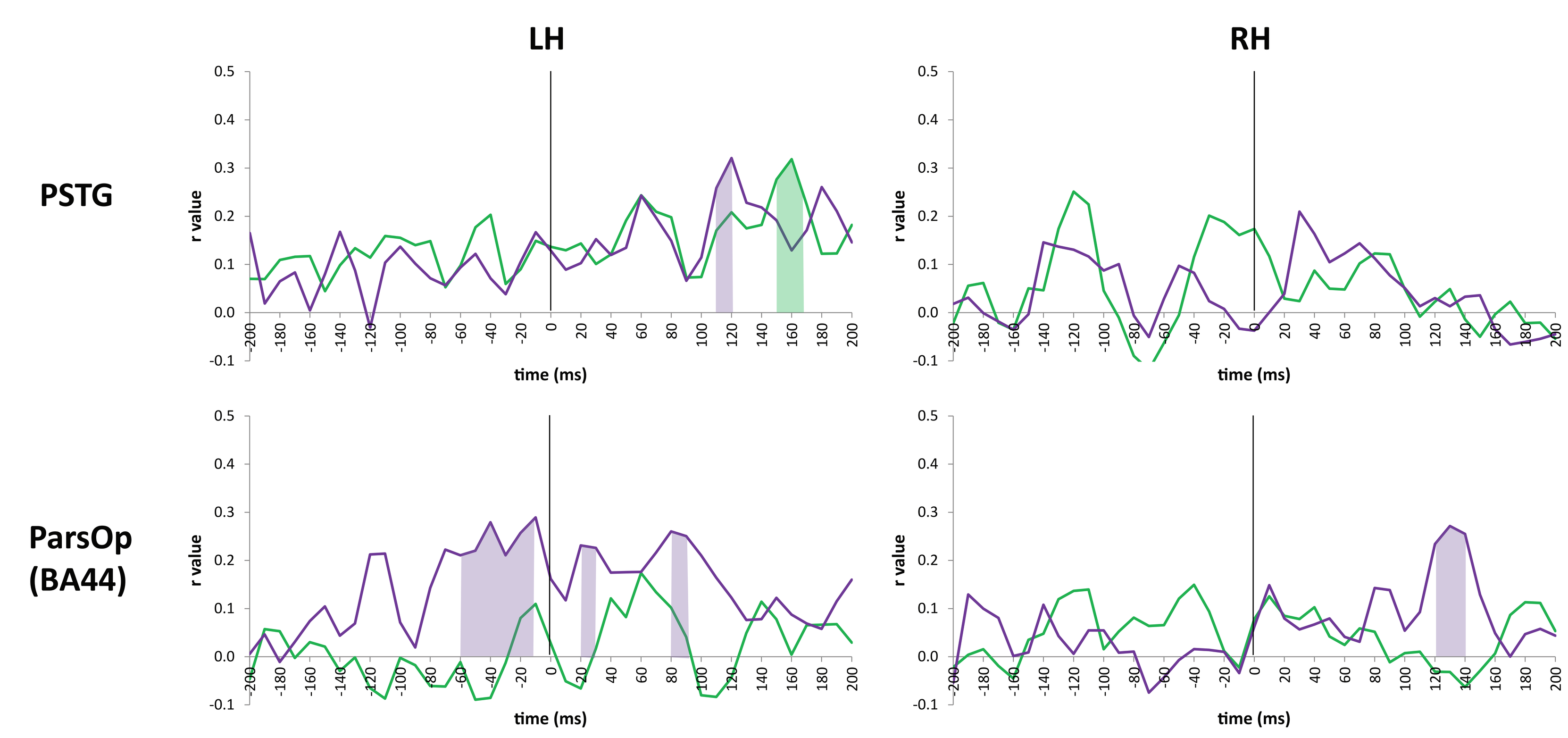


*p < .01

Analyses aligned to affix onset

Thirty regular items (I walk/I jump, He walks/He jumps, He walked/He jumped), averaged over 12 repetitions of each item. Analyses are aligned to the onset of the affix (-s/-ed), shown at 0 ms.





*p < .01

Conclusions

- First dissociation to emerge was between pronouns (I vs. He), and between words and musical rain at 80 ms in bilateral fronto-temporal regions
- Left inferior frontal cortex (pars opercularis) showed sensitivity to differences between affixed and non-affixed words in the time window around affix onset
- Acoustic baseline conditions did not show same distinction, suggesting the affix-related effect cannot be attributed to length difference or other low-level features
- RSA results suggest speech-specificity emerges by 80 ms and is modulated by acoustic properties of the input (differences between I and He)
- Processing within left inferior frontal cortex is sensitive to the presence of an affix, consistent with neuroimaging results showing this region plays a key role in morphological processing
- Multivariate RSA technique allows inferences about qualitative properties of the underlying processes over space and time

References

[1] Marslen-Wilson WD & Tyler LK (2007) *Phil Trans R Soc B, 362*, 823-836.
[2] Bozic M, Tyler LK, Ives DT, Randall B, & Marslen-Wilson WD (2010) *Proc Natl Acad Sci*, *107*(*45*), 17439-17444.
[3] Kriegeskorte N, Mur M, & Bandettini PA (2008) *Frontiers in System Neuroscience*, *2*, 1-27.
[4] Su L, Fonteneau E, Marslen-Wilson WD & Kriegeskorte N (2012) *PRNI 2012.*

Contact **cmw59@cam.ac.uk**