Tracking the neural dynamics of speech comprehension using representational similarity analysis (RSA)

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

Introduction

Extensive neuroimaging and neuropsychological evidence has revealed a bilateral fronto-temporal network supporting speech comprehension (Marslen-Wilson & Tyler, 2007) and a left-lateralised sub-system specialised for processing of grammatical complexity (e.g. *walk-ed*; Bozic et al., 2010).

Using combined MEG-EEG, we addressed how neural activity in these networks is modulated by properties of the speech input. We manipulated the presence of an inflectional affix (*–s* and past-tense *–ed*) to investigate what regions are sensitive to morphological complexity.

The application of multivariate representational similarity analysis (RSA; Kriegeskorte et al., 2008) allowed us to examine the information carried by patterns of activity in a region at a particular time, rather than overall activation.

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, which was derived by jittering formants in each sound file. Musical Rain 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 previous verb phrase was coherent within a sentence (e.g. *He walked* "to the store").

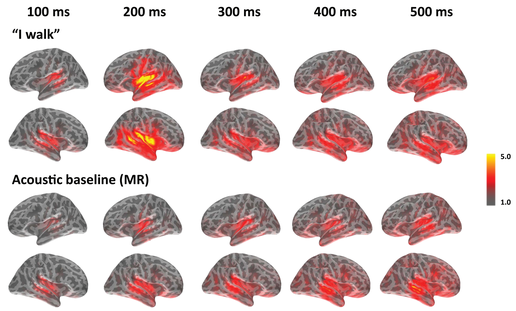
Acquisition and Source Reconstruction

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.

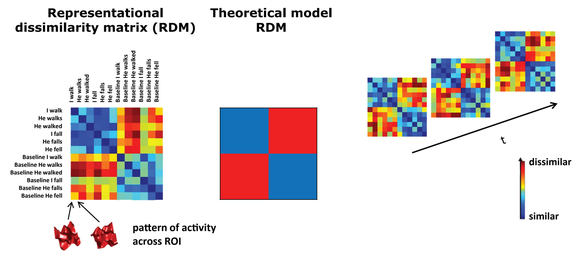
Results

Source level (noise-normalised MNE)



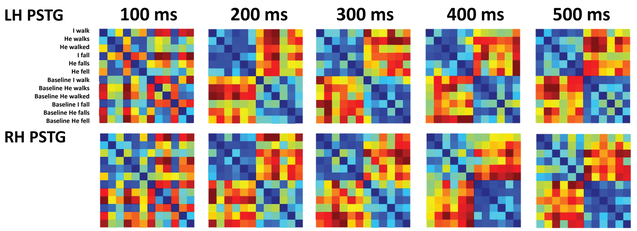
Mean source activity for non-affixed condition (I walk) and matched Musical Rain from 100 to 500 ms. Increased activation appeared in speech-related areas for words compared to the Musical Rain baseline.

Representational Similarity Analysis (RSA)



Patterns of neural 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 that are emphasised in a given region. The further dimension of time provides us with one matrix at every time point.

Data RDMs

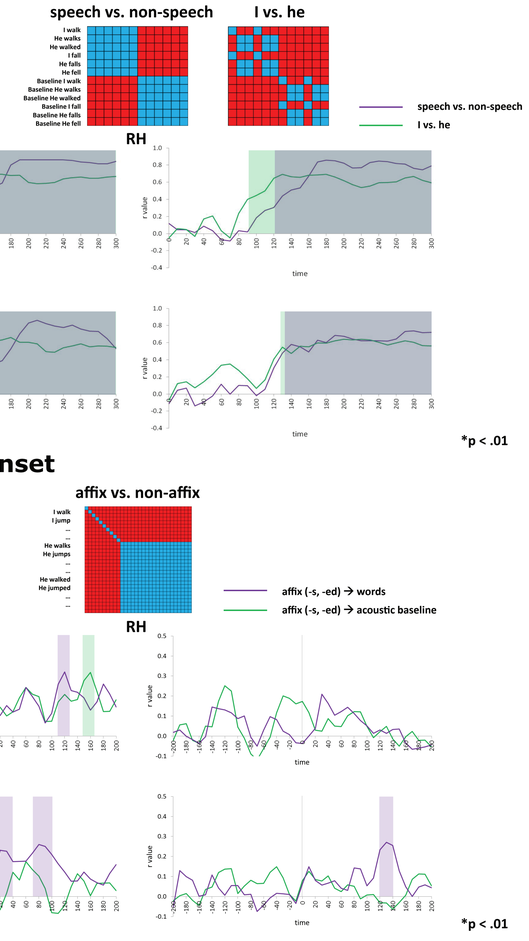


Mean activation RDMs for six word and six Musical Rain conditions in bilateral posterior superior temporal gyrus (PSTG).

Results

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

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 is 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) is the only region to show sensitivity to differences between affixed and non-affixed words from 50 ms before to 100 ms after affix-onset
- Acoustic baseline conditions does not show same distinction
- 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 BA44 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

Marslen-Wilson, W.D. & Tyler, L.K. (2007). Morphology, language and the brain: the decompositional substrate for language comprehension. *Phil. Trans. R. Soc. B, 362*, 823-836.
Bozic, M., Tyler, L.K., Ives, D.T., Randall, B., & Marslen-Wilson, W.D. (2010). Bihemispheric foundations for human speech comprehension. *Proceedings of the National Academy of Sciences*, *107*(45), 17439-17444.
Kriegeskorte, N., Mur, M., & Bandettini, P.A. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in System Neuroscience*, 2, 1-27.

cmw59@.cam.ac.uk