Transforming acoustic input into a hierarchy of linguistic representations

leg5@nyu.edu
@GwilliamsL

Laura Gwilliams[1,2] David Poeppel[1,3] Jean-Remi King[1,4]

New York University[1]; NYU Abu Dhabi[2]; Max Planck Institute[3]; CNRS[4]

1 Questions

- To comprehend continous speech, the brain needs to generate a hierarchy of linguistic representations
- It is currently unknown which are the fundamental representations and how they are orchestrated

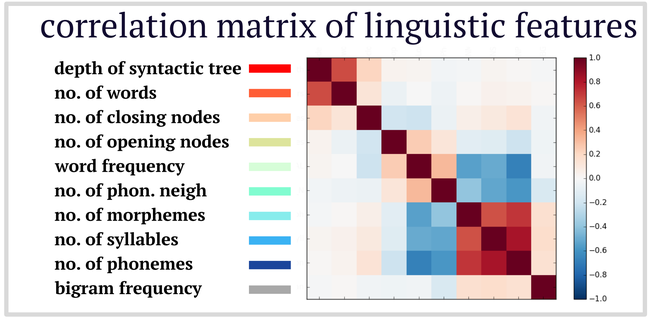
a. **which linguistic units are encoded in brain responses to naturalistic speech?**

b. **what is the relative time-course with which these properties are processed?**

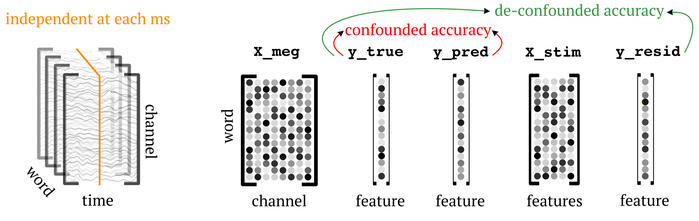
c. **what computational architecture supports these linguistic representations?**

2 Method

- 18 participants listened to four stories (twice)
- 2 x one hour recordings
- KIT 208 channel MEG system
- Comprehension questions every ~4 minutes
- Responses to ~8000 words per participant

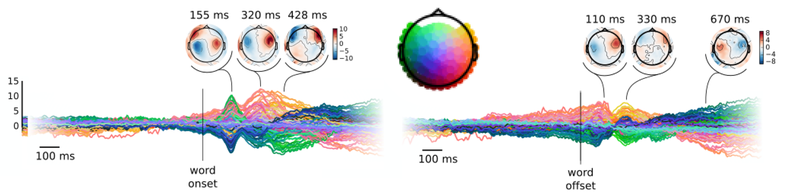


- Stories were fully annotated - from acoustic to lexical and syntactic properties

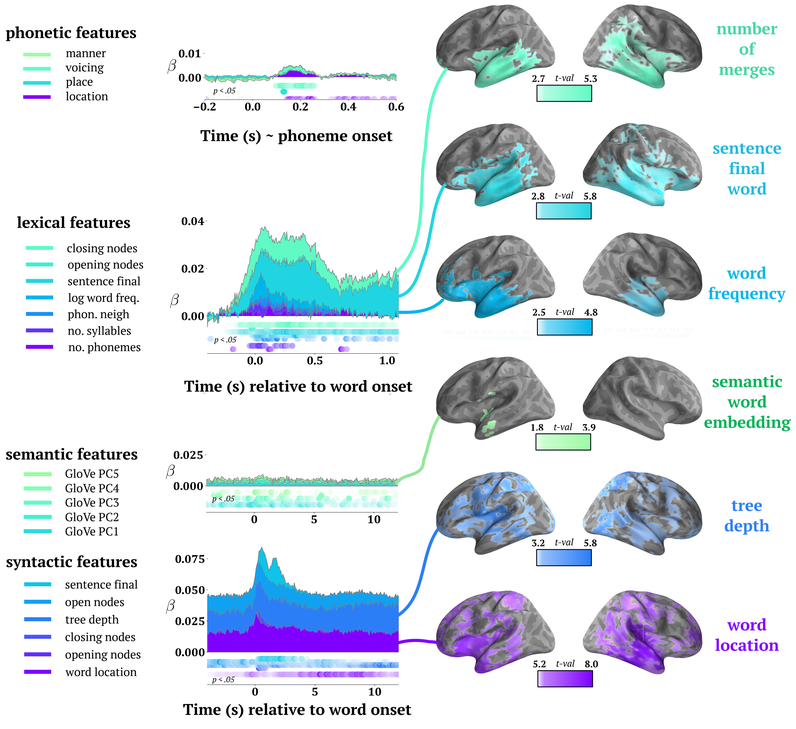


- Due to the high natural co-variation between different linguistic properties, we need to orthogonalise them

3 Decoding the Hierarchy

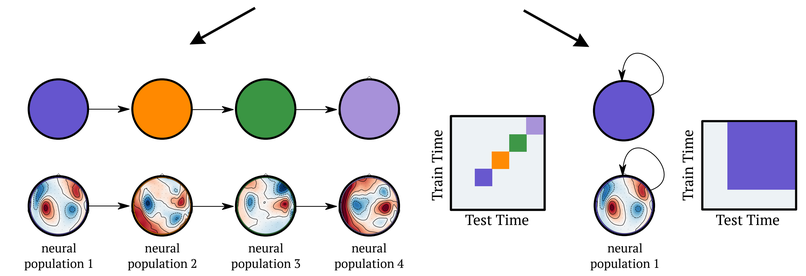


- Apply algorithm at each time-sample to derive a timecourse of decoding accuracy, and at each spatial location to derive a spatial map of decoding accuracy
- Analysis is time-locked to different sized units (phonemes, words) across different time-scales (sub-second, second and tens of seconds)



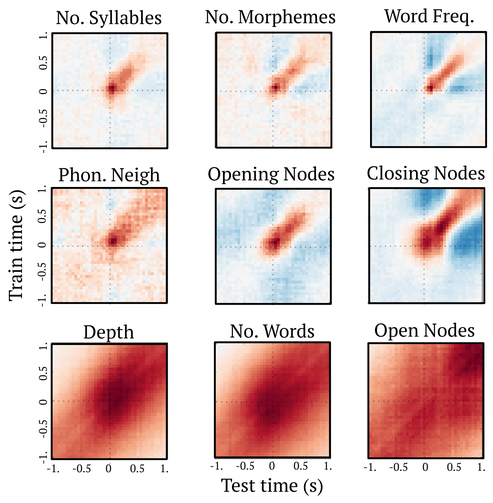
- Can decode a rich set of representations from the MEG signal, spanning the entire hierarchy from lexical to syntactic properties
- Significant temporal overlap between representations
- Evidence for some contentious units: syllables and morphemes
- Some representations are specifically locked to word onset/offset, whereas others are sustained over time

4 Feedforward or Recurrent?



Dichotomy between processes occurring on a local scale, where each region is active for 100-200 ms, and processes which occur on a more global scale, where each region is active for 1-2 seconds.

A combination of both feedforward and recurrent processes are recruited depending on the linguistic representation



5 Conclusion

- The brain generates multiple representations concurrently, spanning the entire linguistic hierarchy
- We find evidence for some contentious units (no. of morphemes; no. of syllables)
- Both feedforward and recurrent computations are recruited, depending on the linguistic representation

Three key aspects of the neural architecture supporting speech comprehension: parallel (3), feedforward and recurrent (4) computations

Funding: NYU Abu Dhabi G1001; EU 660086; NIH 2R01DC05660