



HAL
open science

How distinct are Syntactic and Semantic Representations in the Brain During Sentence Comprehension?

Subba Reddy Oota, Frédéric Alexandre, Xavier Hinaut

► **To cite this version:**

Subba Reddy Oota, Frédéric Alexandre, Xavier Hinaut. How distinct are Syntactic and Semantic Representations in the Brain During Sentence Comprehension?. SNL 2022 - 14th Annual Meeting of the Society for Neurobiology of Language, Oct 2022, Philadelphia, United States. 2022. hal-03942241

HAL Id: hal-03942241

<https://inria.hal.science/hal-03942241v1>

Submitted on 16 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Introduction

- 1 Syntactic parsing is the task of assigning a syntactic structure to a sentence.
- 2 Recent works have used syntactic embeddings from constituency trees and other word syntactic features to understand how syntax structure is represented in the brain's language network.
- 3 However, the effectiveness of dependency parse trees or the relative predictive power of the three syntax parsers is yet unexplored.

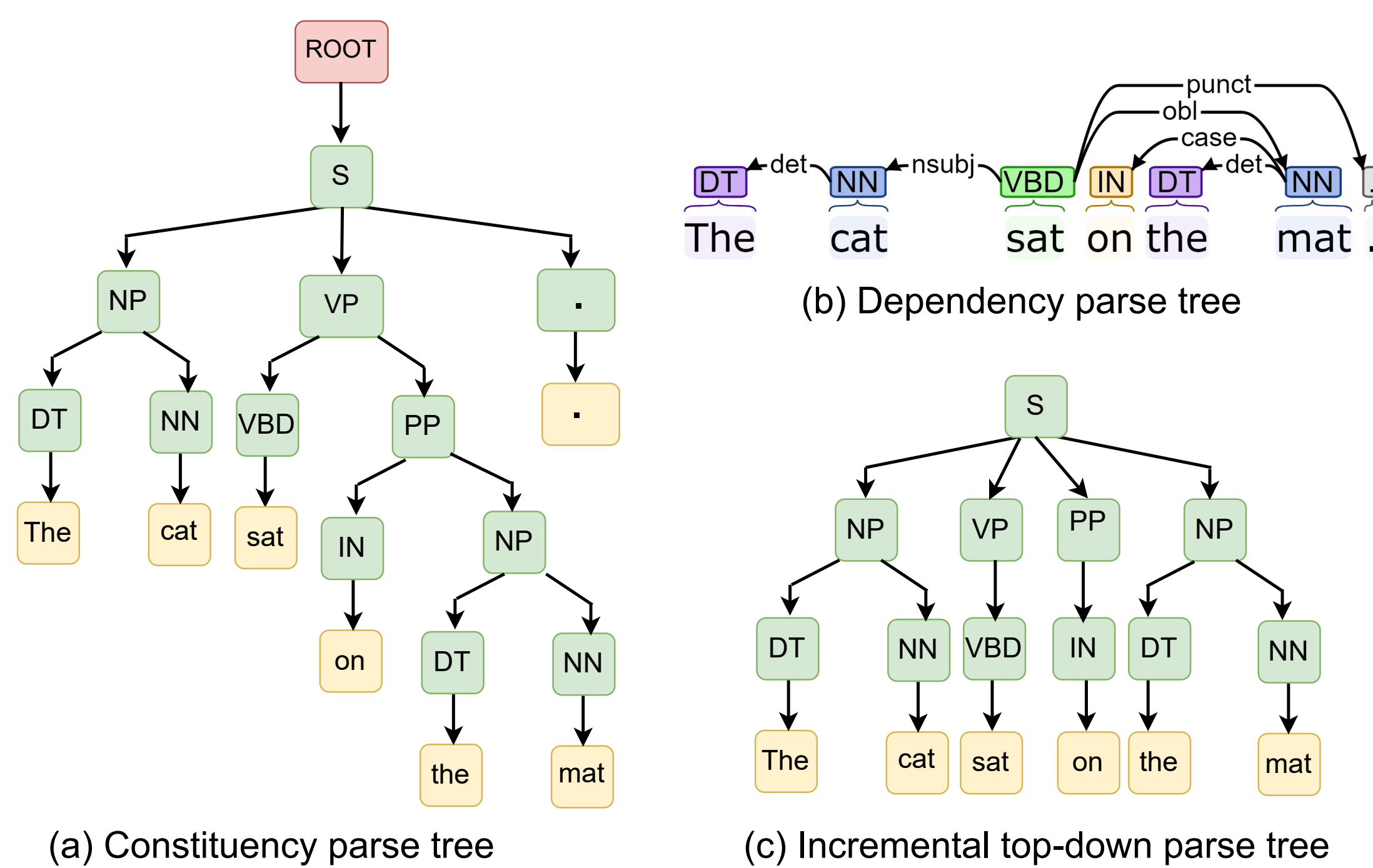


Figure 1: Example of syntactic parsing for three parsers

Questions

- 1 Do syntactic parser representations predict similar fMRI areas?
- 2 Which syntactic parsing method better predicts fMRI activity?

Main Contributions

- 1 We explore syntactic structure embeddings obtained from three parsers and use them in an encoding model to predict brain responses.
- 2 We use a GCN model (SynGCN embeddings) for the dependency parser that accurately encodes the global syntactic information.

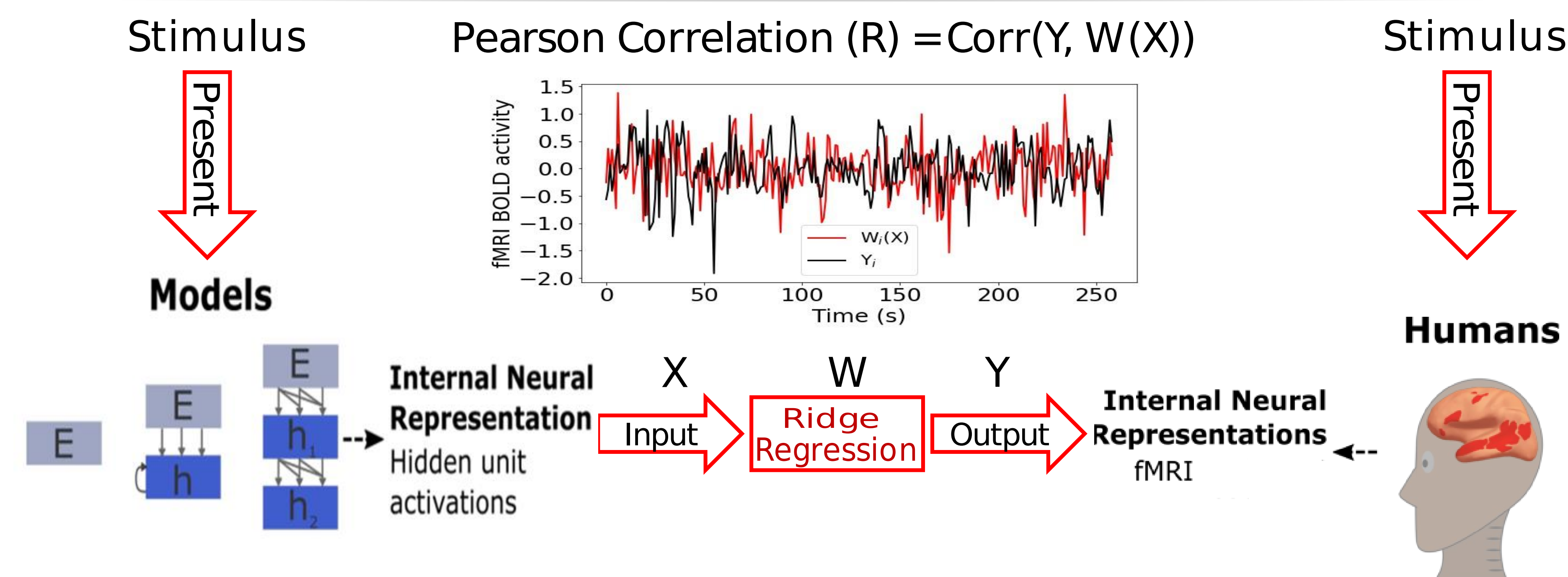
Feature Representations

- 1 Constituency Tree-based Embeddings
- 2 Dependency Tree-based Embeddings using GCN (SynGCN)
- 3 Incremental Top-Down Parser Embeddings
- 4 Basic Syntactic Features
- 5 BERT Features

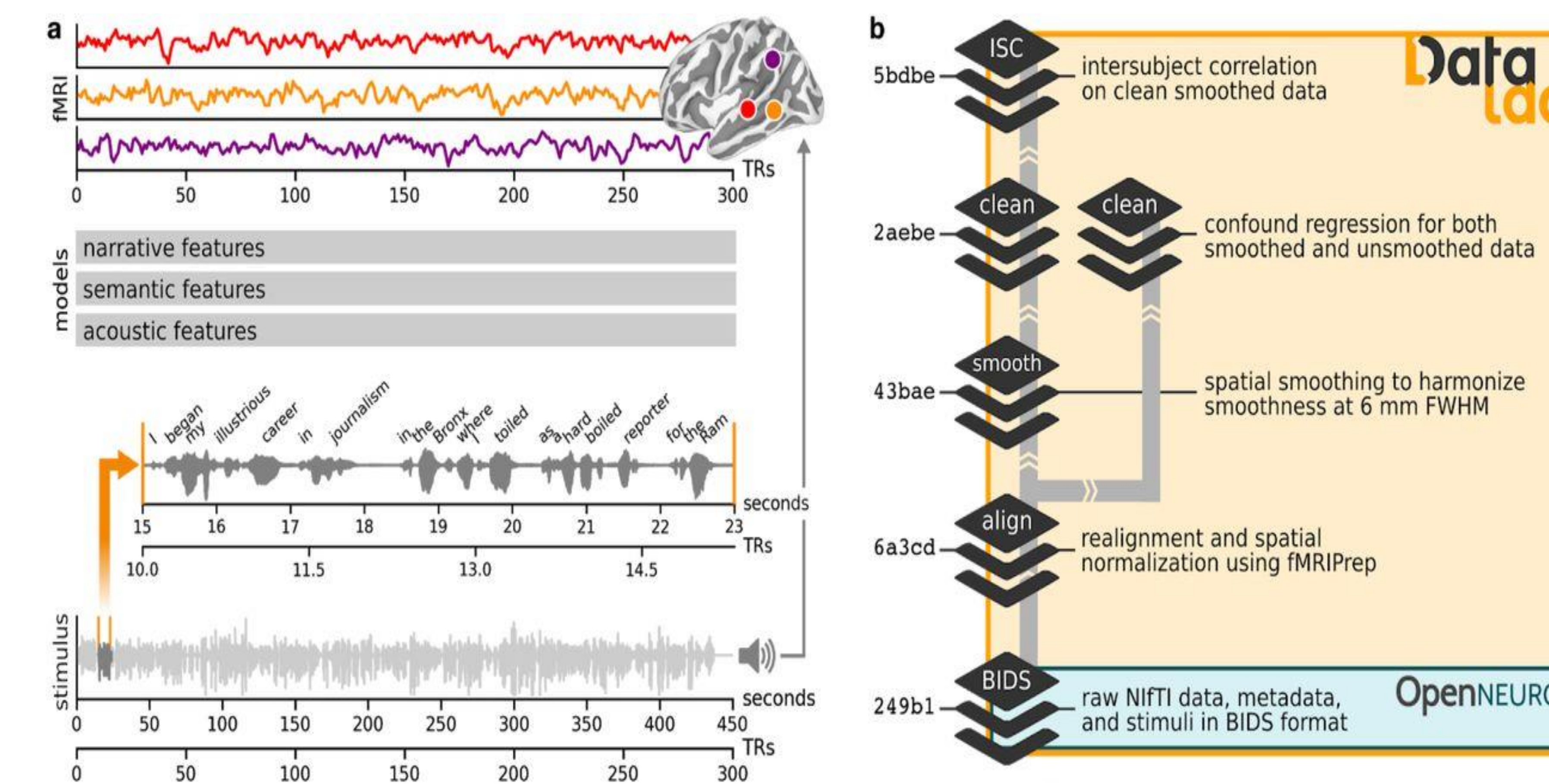
Evaluation Metrics

- 1 R^2 Score
- 2 Pearson Correlation (R)

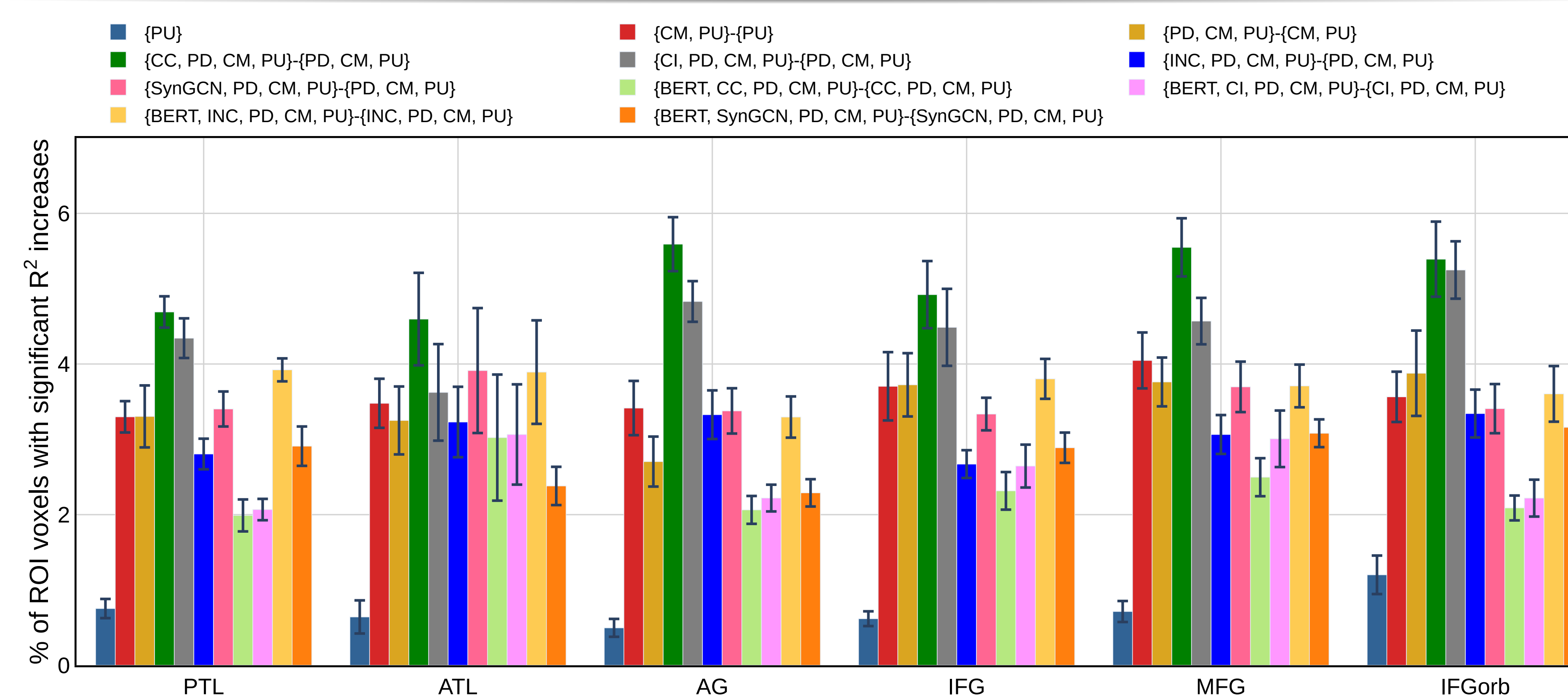
Brain Encoding



Narrative Listening (Pieman dataset from Nastase et al. 2021)



Encoding Performance of Various Feature Sets



Brain Maps (Pearson Correlation)

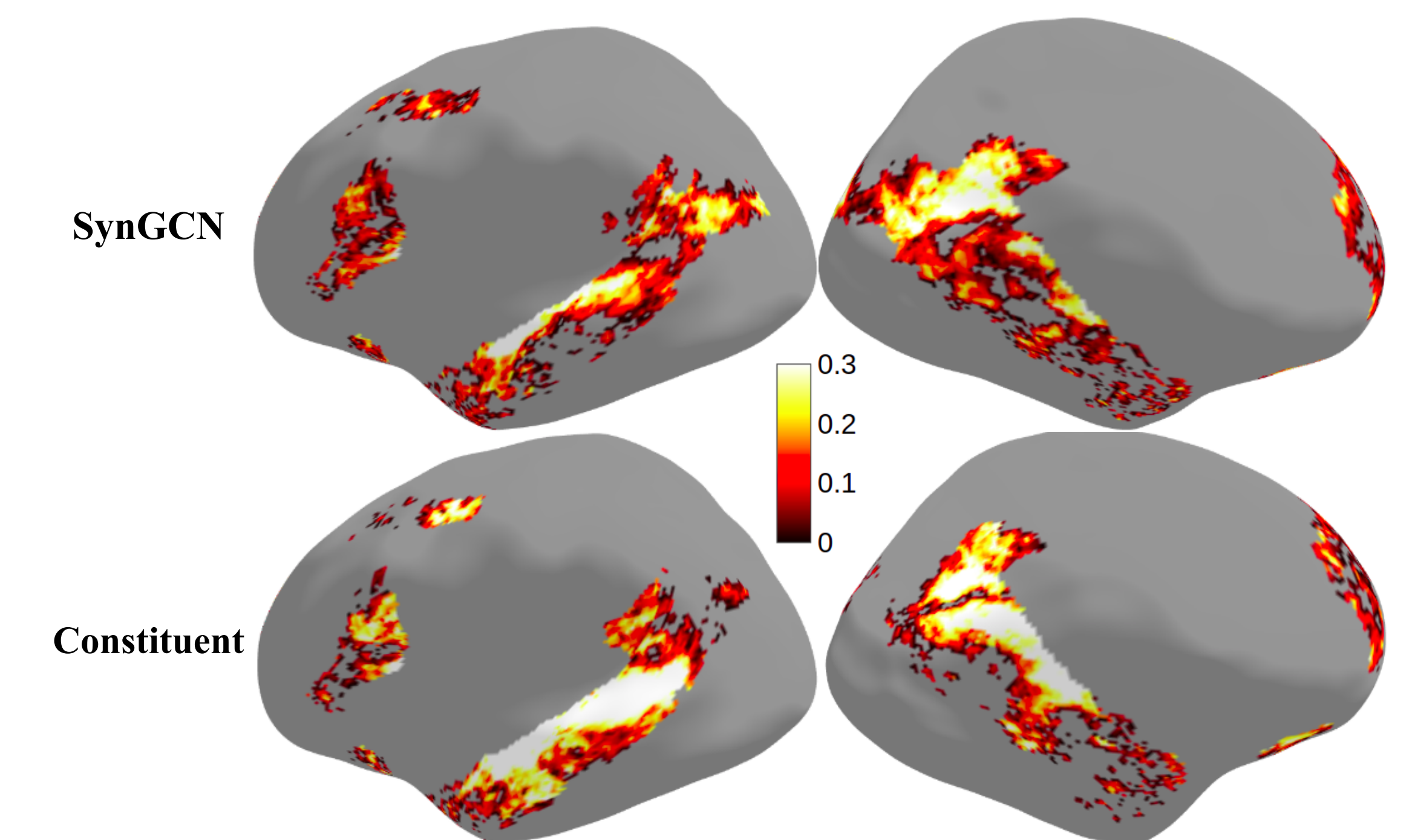


Figure 2: Whole Brain

Cognitive Insights

- 1 PTL and ATL have high overlap for both syntactic (CC, CI, SynGCN, INC), and semantic (BERT) features.
- 2 ROIs such as IFG, IFGorb and MFG have higher overlap with syntactic features like CM, PU, SynGCN and INC, and with semantic feature BERT.
- 3 Basic syntactic features are much less associated with voxels in angular gyrus region.

Conclusion

- 1 Constituency trees explain additional variance better than other syntactic parsing methods.
- 2 Future Directions: This work was done on data related to English stories only. As we do other kinds of models of language processing in various languages, we want to make similar studies for multi-lingual languages [1].

References

- [1] Bhattasali Li Jixing et al. 2022. Le petit prince multilingual naturalistic fmri corpus.
- [2] Aniketh Janardhan Reddy et al. 2021. Can fmri reveal the representation of syntactic structure in the brain?
- [3] Samuel A Nastase et al. 2021. The "narratives" fmri dataset for evaluating models of naturalistic language comprehension.

Acknowledgements

- 1 We thank Gaël Jobard for his very helpful comments and discussions.
- 2 We thank Inria for the PhD funding and ANR DeepPool project.