Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] They are described in the Discussion and Related Work section.
   (c) Did you discuss any potential negative societal impacts of your work? [N/A] This paper makes purely scientific contributions. Thus, we do not see any direct societal impacts because of this work.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] They are available at https://github.com/anikethjr/brain_syntactic_representations.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] We mention them in the methods section wherever appropriate.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] They are shown wherever appropriate. Figure 4 that highlights ROI-level prediction performance shows error bars. More detailed plots for each ROI are in Appendix D. These plots also show data points for each individual subject.
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] A rough estimate and a brief description of the resources used are included in our code's README file.
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes]
   (b) Did you mention the license of the assets? [Yes] They are mentioned in Appendix B.
   (c) Did you include any new assets either in the supplemental material or as a URL? [Yes]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes] This is mentioned in the "fMRI data" subsection of the methods. Wehbe et al. [17]'s protocol was approved by the Carnegie Mellon University Institutional Review Board as mentioned in the methods section
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A] We do not use any data that contains personally identifiable information or offensive content. Wehbe et al. [17]'s protocol was approved by the Carnegie Mellon University Institutional Review Board as mentioned in the methods section.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A] We use the data collected by Wehbe et al. [17]. Please refer to their manuscript for more details about the task.
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A] Wehbe et al. [17]'s protocol was approved by the Carnegie Mellon University Institutional Review Board as mentioned in the methods section. Please refer to their manuscript for more details.
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A] It was not included in the study description [17].

Appendix

A Acquiring and preprocessing the fMRI data

We obtained the raw data from Wehbe et al. [17]. This fMRI data is acquired at a rate of 2s per image and comprise $3 \times 3 \times 3mm$ voxels. The data for each subject is slice-time and motion corrected using SPM8 [57], then detrended and smoothed with an isotropic spherical Gaussian kernel with a standard deviation of $3mm$. The brain surface of each subject is reconstructed using Freesurfer [58] and a grey matter mask is obtained. Pycortex [59] is used to handle and plot the data. All subject results are converted to MNI space using pycortex.

B Licenses of Assets Used

The data we use was published by Wehbe et al. [17]. It is free to use with no restrictions.

We use the code published by Adhikari et al. [31]. The license for this code states that it can be used for educational and research purposes without a fee. However, any for-profit use requires written consent of the authors.

We also use the incremental top-down parser published by Roark [34]. It was published using the Apache License, Version 2.0.

C Raw prediction results

Figure 5 shows the prediction results obtained using each feature group. To be able to better judge different levels of accuracy, instead of looking at the R^2 scores, we compute R^{2+}, in which we replace the positive R^2 values by their squared root, making them easier to resolve visually, and the negative ones with 0. Figure 6 shows the difference in R^{2+} scores between the pairs of feature groups that were tested for significant differences in R^2 scores to generate the results presented in Figures 4 and 3.

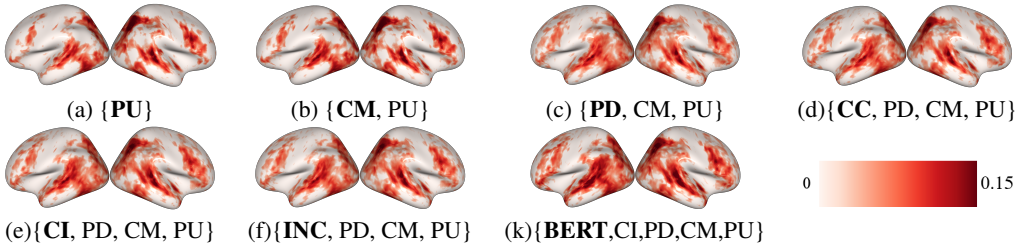

Figure 5: Cross-subject prediction performance of all syntactic feature groups. The figures show cross-subject average R^{2+} scores. Here, PU = Punctuation, CM = All complexity metrics, PD = POS and DEP Tags, CC = ConTreGE Comp, C = ConTreGE Incomp, INC = InConTreGE, BERT = BERT embeddings and '{,}' indicates that these features were concatenated in order to make the predictions.

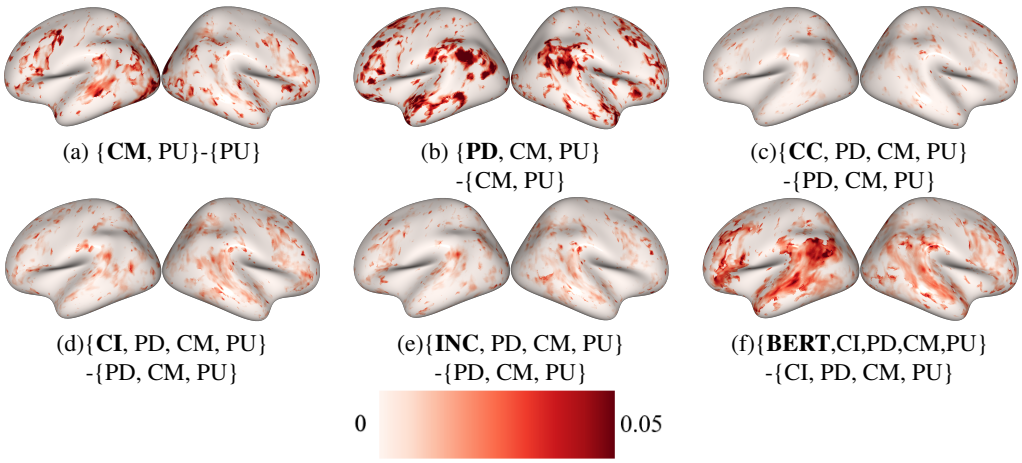

Figure 6: Cross-subject differences in prediction performance between the pairs of feature groups shown in Figures 3 and 4. The figures show cross-subject average differences in R^{2+} scores. Here, PU = Punctuation, CM = All complexity metrics, PD = POS and DEP Tags, CC = ConTreGE Comp, C = ConTreGE Incomp, INC = InConTreGE, BERT = BERT embeddings and '{,}' indicates that these features were concatenated in order to make the predictions. '-' indicates that the second feature group's R^{2+} scores were subtracted from the first feature group's R^{2+} scores.

D ROI Analysis Plots

In this section, we show detailed plots from the ROI analysis of prediction performance. Each plot corresponds to one ROI from Figure 4. Like in Figure 4, each bar represents the average percentage of significantly predicted ROI voxels across subjects and the error bars show the standard error across subjects. The labels for the bars denote the feature that was added while performing the significance test. Thus, the labels used here correspond to the following labels used in Figures 3 and 4:

1. CM = {CM, PU} - {PU}
2. PD = {PD, CM, PU} - {CM, PU}
3. CC = {CC, PD, CM, PU} - {PD, CM, PU}
4. CI = {CI, PD, CM, PU} - {PD, CM, PU}
5. INC = {INC, PD, CM, PU} - {PD, CM, PU}
6. BERT = {BERT, CI, PD, CM, PU} - {CI, PD, CM, PU}

Each red point then shows the percentage of ROI voxels that were significantly predicted, according to the test specified by the bar label, in a given subject.

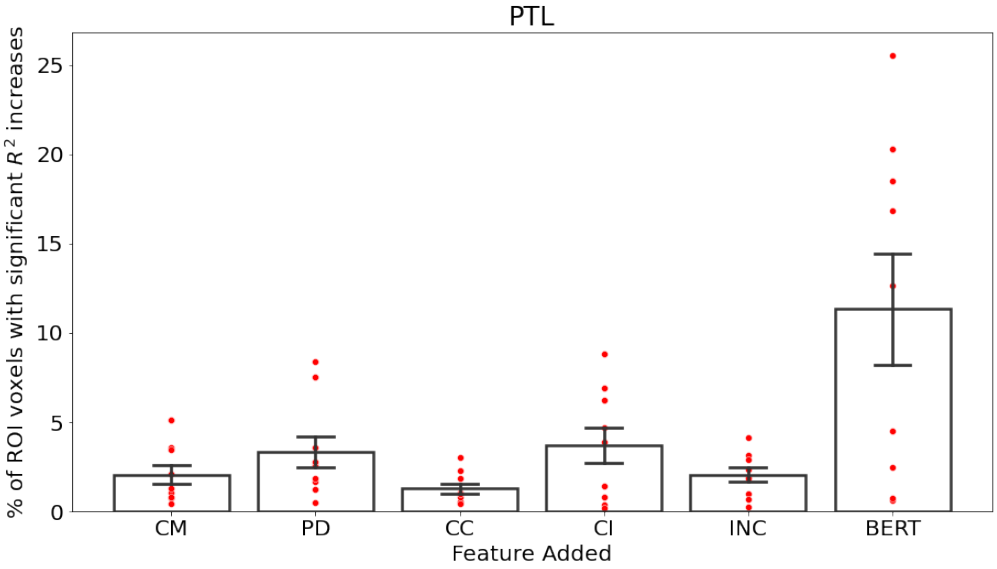

Figure 7: Results for the Posterior Temporal Lobe

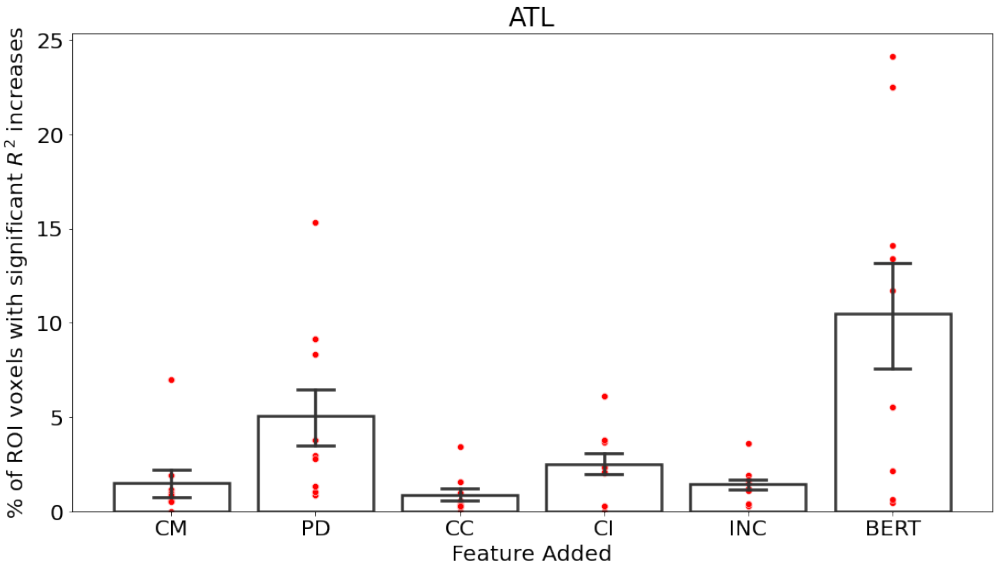

Figure 8: Results for the Anterior Temporal Lobe

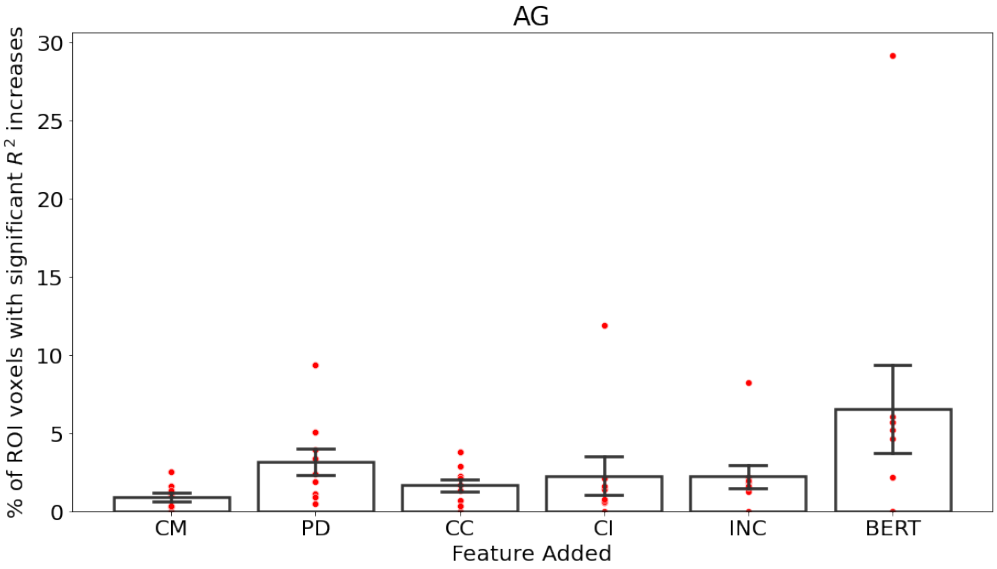

Figure 9: Results for the Angular Gyrus

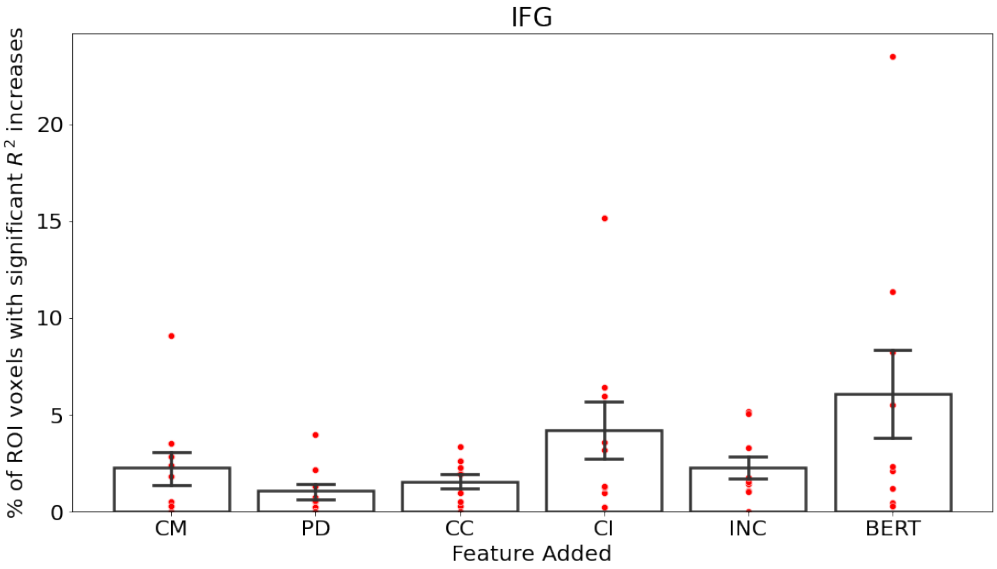

Figure 10: Results for the Inferior Frontal Gyrus

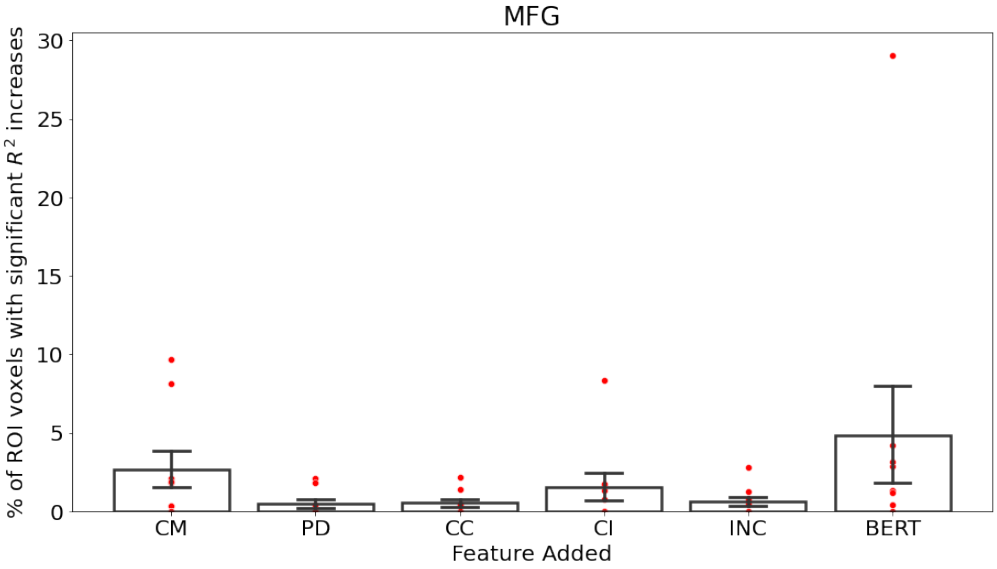

Figure 11: Results for the Middle Frontal Gyrus

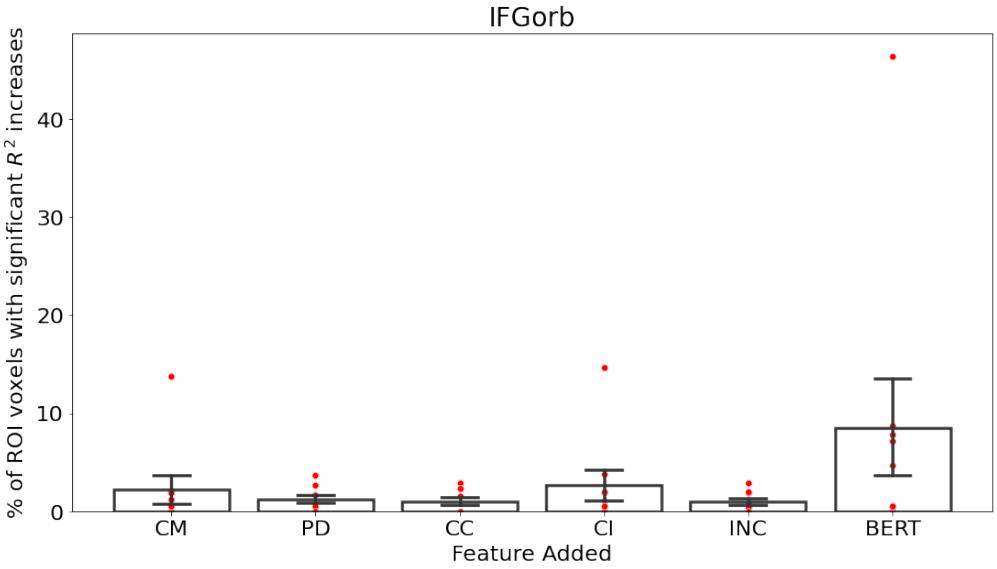

Figure 12: Results for the Inferior Frontal Gyrus Orbital