Left Hemisphere Neural Dynamics during Spontaneous vs. Cued Speech Production

Nikita Agrawal[1,2], Werner Doyle[2,3], Orrin Devinsky[1,2], Adeen Flinker[1,2]

Department of Neurology[1], Comprehensive Epilepsy Center[2], Department of Neurosurgery[3], NYU School of Medicine

NYU Langone Health

Introduction

A variety of speech production tasks are used to localize language for surgical planning to avoid postoperative language deficits. Neuroimaging studies in fMRI and PET have shown that overlearned speech production, such as number counting, does not reliably activate left hemisphere language cortex [1,2]. Similarly, electrical stimulation of cortex during counting does not reliably produce a speech deficit [3]. While previous electrocorticography (ECoG) studies have mainly focused on cued speech task production and have linked left inferior frontal gyrus activation with the pre-articulatory phase of speech production, the timing and degree of left hemisphere recruitment during spontaneous speech remains underspecified [4,5]. **Here, we model high gamma ECoG data to examine the spatial and temporal dynamics of several spontaneous and cued speech production tasks.**
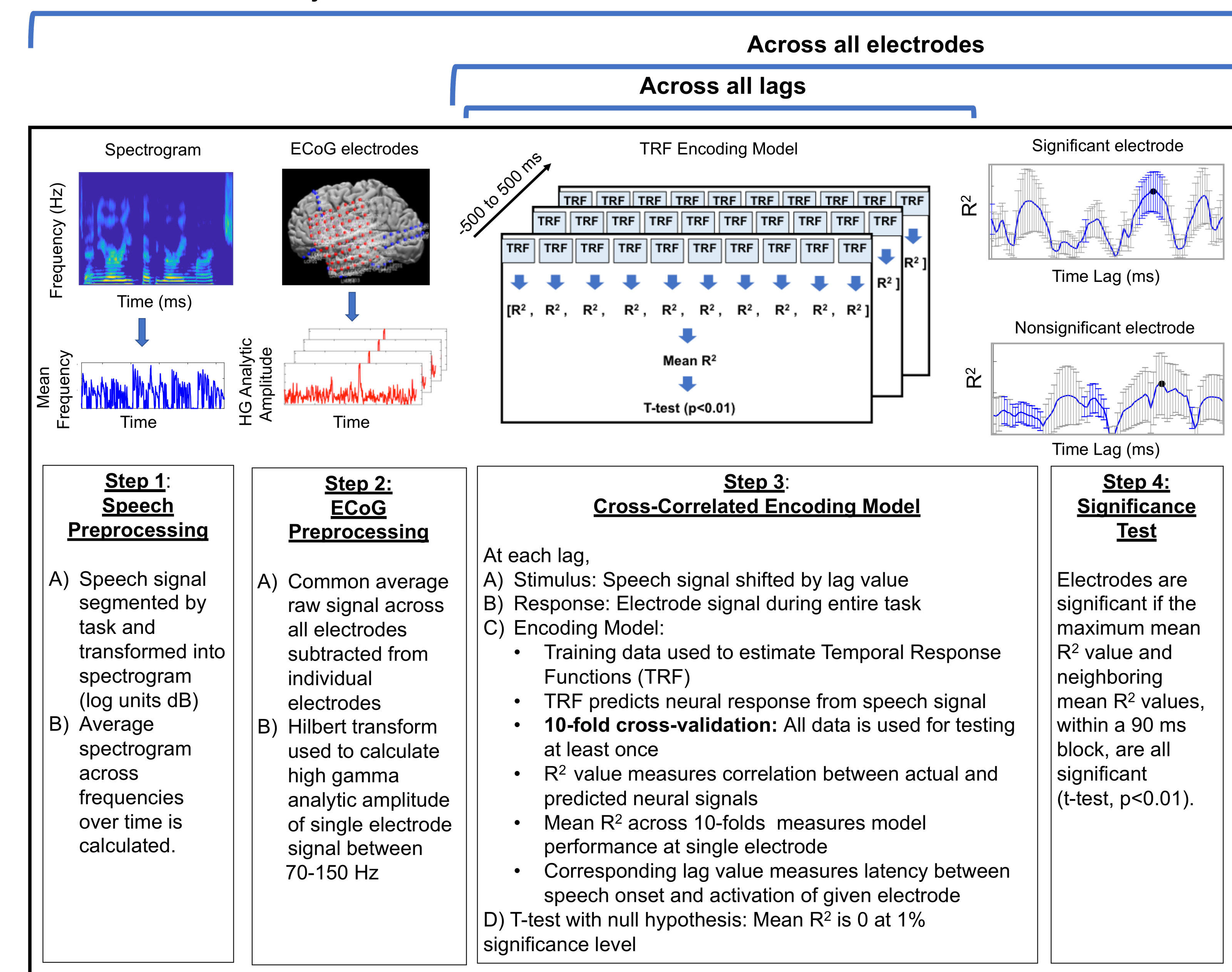
Methods

Subjects: 4 consenting patients with intractable epilepsy, implanted with ECoG grids of up to 256 electrodes, in left hemisphere language cortex.
Tasks: Microphone records the subject's voice in parallel with the ECoG recording during 5 tasks:

Task	Type	Example
Number Counting	Spontaneous	1 – 20
Months of the Year Recitation	Spontaneous	January – December
High Frequency Sentence Repetition	Spontaneous	Today is a sunny day in NYC.
Low Frequency Sentence Repetition	Spontaneous	The pastry chef was elated.
Word Reading	Cued	apple, axe, cake, leg, violin

Analysis: We used a 10-fold cross-validated encoding model based on regularized reverse cross-correlation to predict neural activity from speech data.

Across all tasks and subjects
Across all electrodes
Across all lags

Step 1: Speech Preprocessing
A) Speech signal segmented by task and transformed into spectrogram (log units dB)
B) Average spectrogram across frequencies over time is calculated.

Step 2: ECoG Preprocessing
A) Common average raw signal across all electrodes subtracted from individual electrodes
B) Hilbert transform used to calculate high gamma analytic amplitude of single electrode signal between 70-150 Hz

Step 3: Cross-Correlated Encoding Model
At each lag,
A) Stimulus: Speech signal shifted by lag value
B) Response: Electrode signal during entire task
C) Encoding Model:
• Training data used to estimate Temporal Response Functions (TRF)
• TRF predicts neural response from speech signal
• **10-fold cross-validation:** All data is used for testing at least once
• R^2 value measures correlation between actual and predicted neural signals
• Mean R^2 across 10-folds measures model performance at single electrode
• Corresponding lag value measures latency between speech onset and activation of given electrode
D) T-test with null hypothesis: Mean R^2 is 0 at 1% significance level

Step 4: Significance Test
Electrodes are significant if the maximum mean R^2 value and neighboring mean R^2 values, within a 90 ms block, are all significant (t-test, p<0.01).

Results

A) Overall recruitment per task

1. Lower degree of neural recruitment in spontaneous speech compared to cued speech

Counting   Months   HF Sentence   LF Sentence   Visual Reading
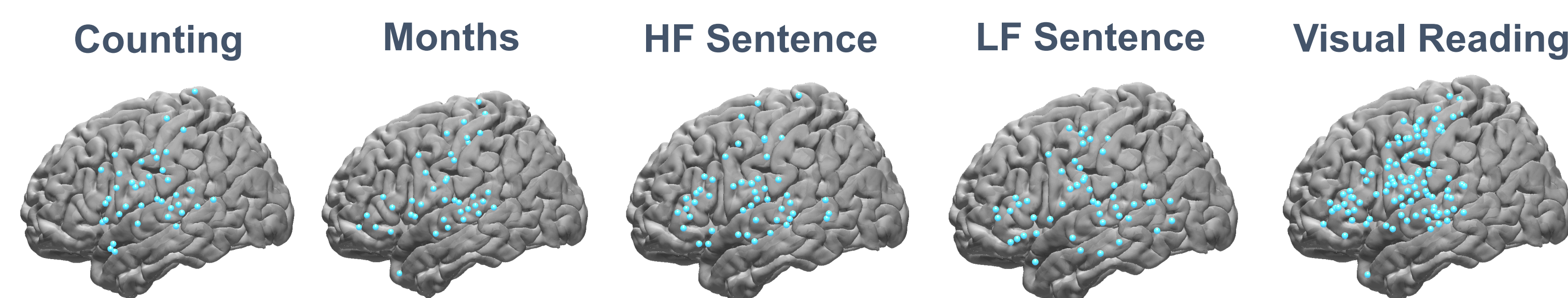
A total of 223 IFG, STG, MTG and peri-central electrodes across all 4 subjects were measured. Only electrodes that passed the significance test were plotted in a normalized space. The number of significant electrodes doubled from spontaneous speech tasks to the cued speech task: 40 (Counting), 40 (Months), 50 (HF Sentence), 51 (LF Sentence) and 104 (Visual Reading).

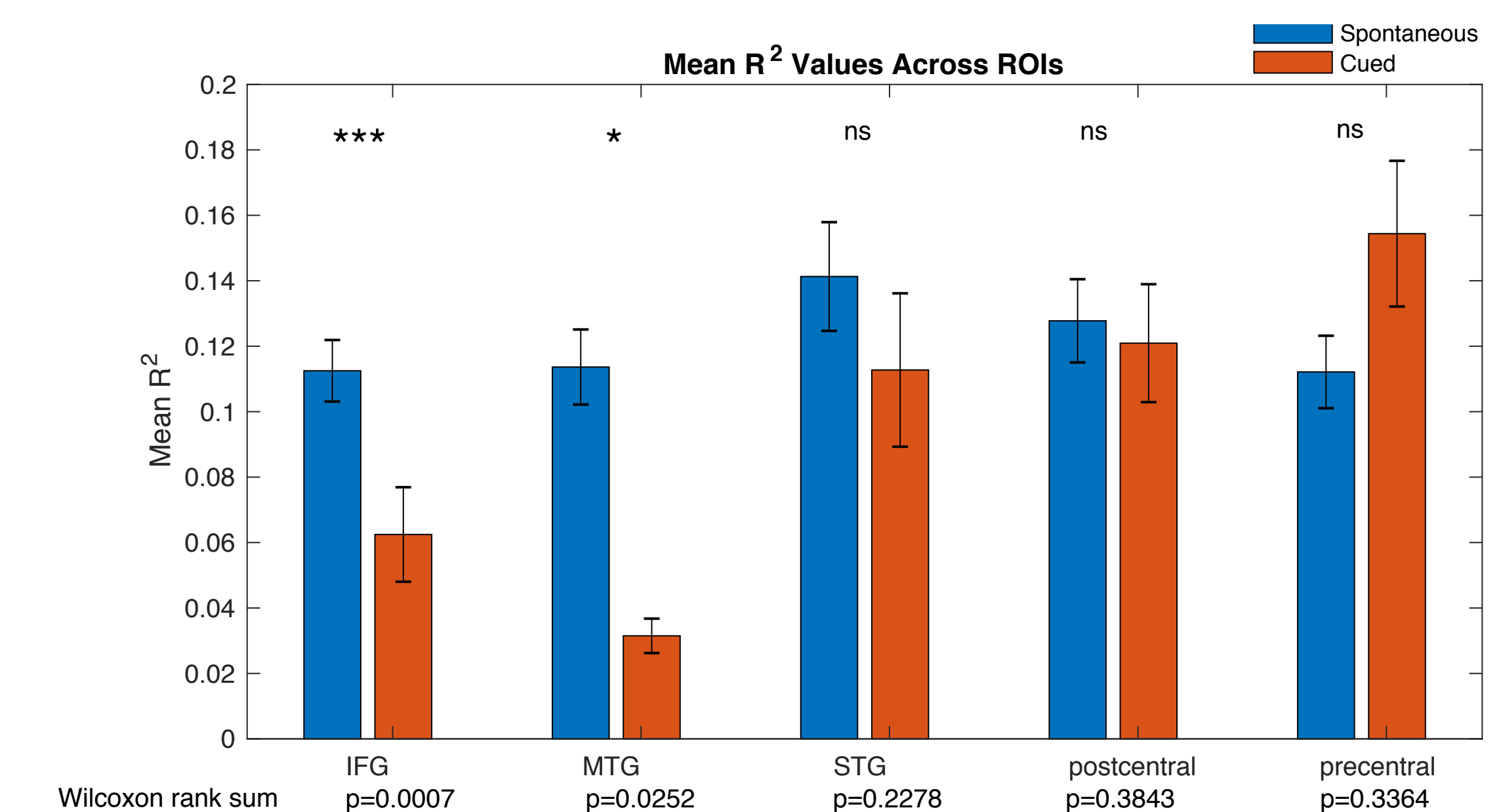### B) Per ROI: Recruitment, average high gamma and mean R^2

2. Within IFG and other regions, graded recruitment as task complexity increases
3. Across ROIs, sentences yield significantly higher mean R^2 than word reading
4. Across ROIs and tasks, mean high gamma does not vary
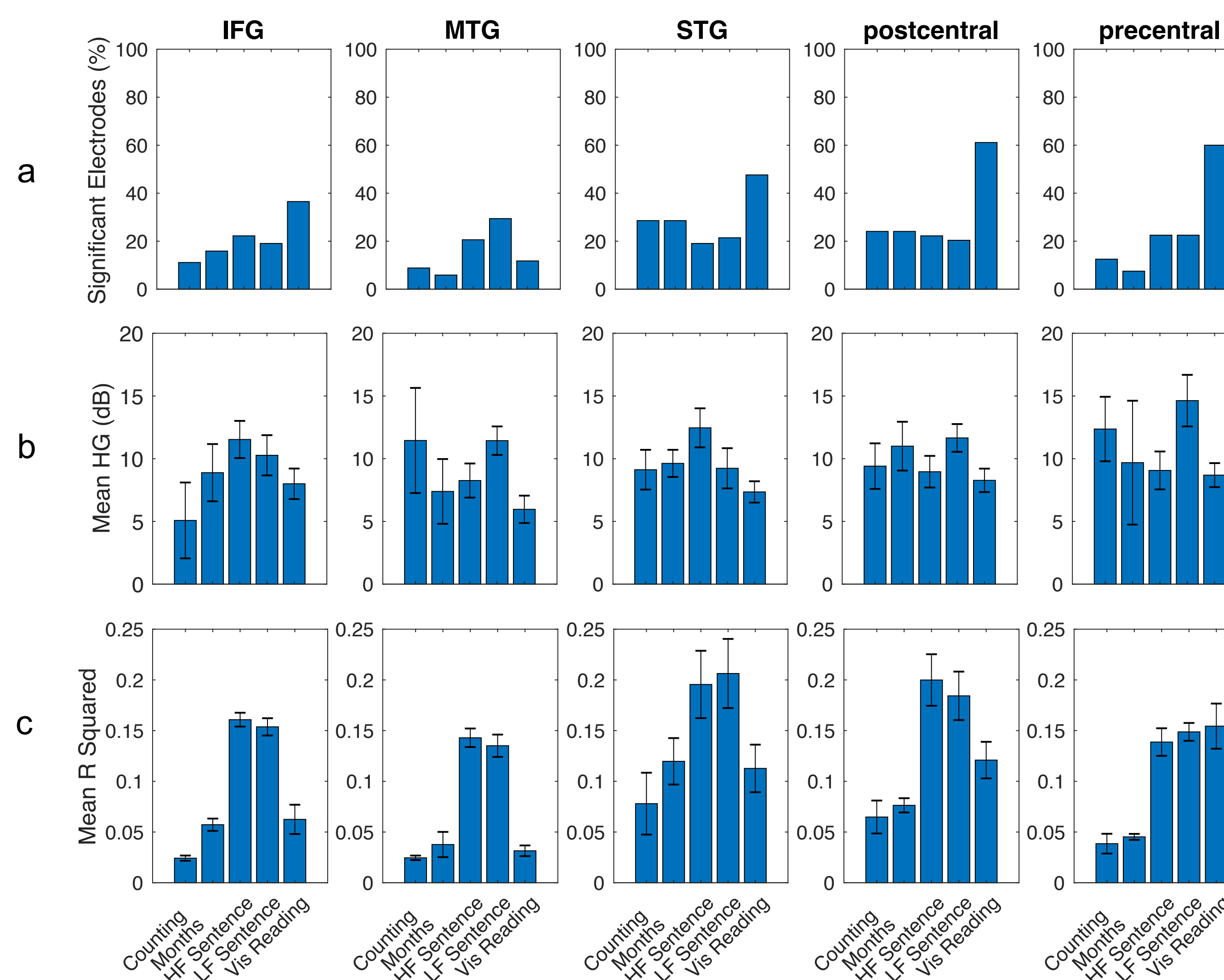
While the cued speech task yields higher recruitment compared to spontaneous speech (a), recitation of high frequency and low frequency sentences consistently yield higher mean R^2 values than word reading, across ROIs (c). In a nonparametric one-way ANOVA (Kruskal-Wallis test), mean R^2 values differ significantly for factor task (p =1.424e-20) but not factor ROI (p=0.1598). High gamma does not vary significantly across tasks and regions, possibly due to averaging across the entire task duration.
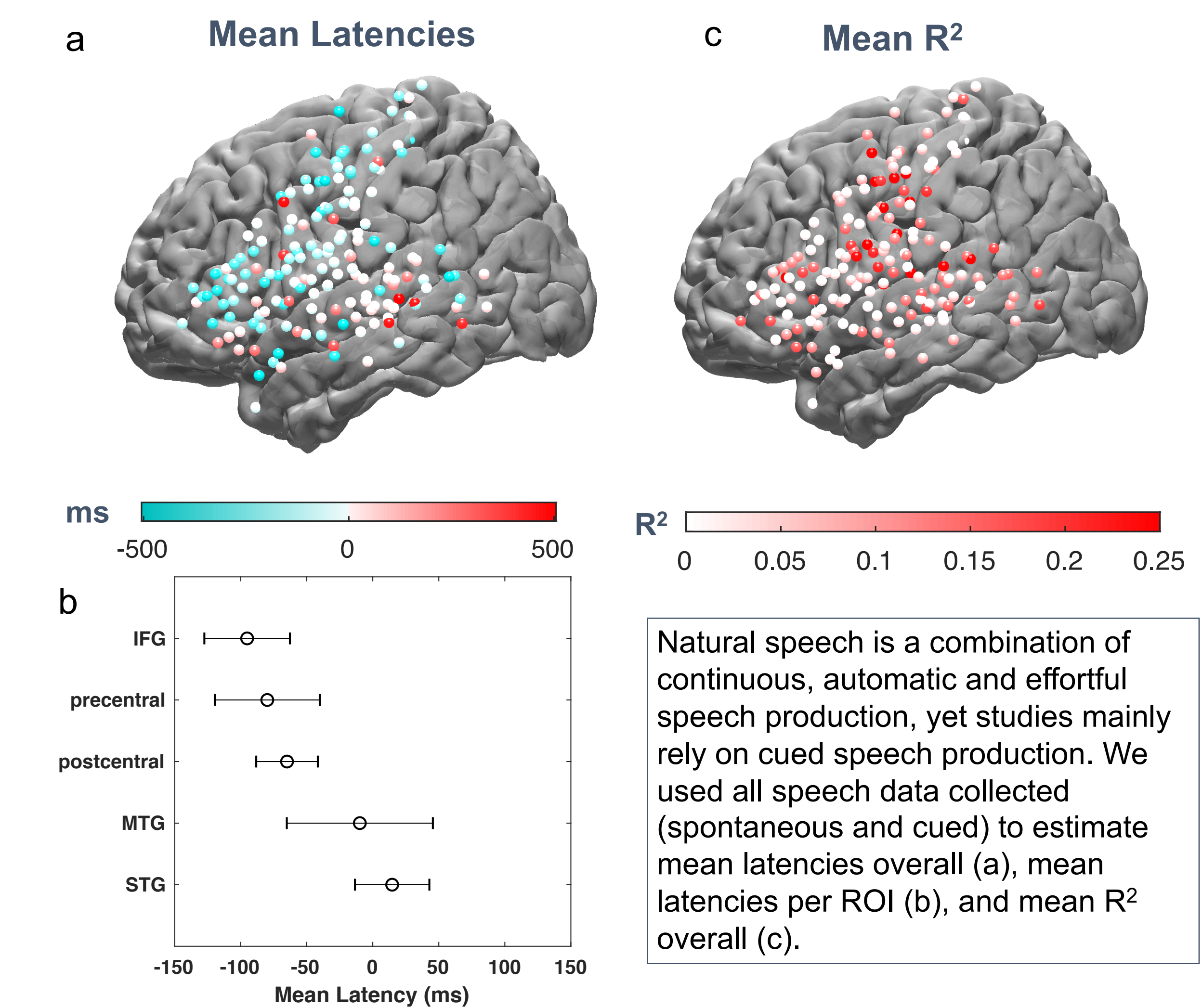
C) Model performance across ROIs

5. Mean R^2 for all Spontaneous tasks vs. Cued task

Mean R^2 Values Across ROIs

Wilcoxon rank sum:
IFG p=0.0007, MTG p=0.0252, STG p=0.2278, postcentral p=0.3843, precentral p=0.3364

Comparison of mean R^2 values across ROIs for combined spontaneous tasks vs. cued tasks indicates whether the model's ability to predict neural activity is enhanced by speech task type. Spontaneous speech yielded significantly higher correlations than cued speech in IFG and MTG. Note, the number of electrodes per ROI varied.

D) Estimating left hemisphere timing and distribution

6. Combining all data for a more realistic speech dataset

a Mean Latencies   c Mean R^2

Natural speech is a combination of continuous, automatic and effortful speech production, yet studies mainly rely on cued speech production. We used all speech data collected (spontaneous and cued) to estimate mean latencies overall (a), mean latencies per ROI (b), and mean R^2 overall (c).

Conclusion

• **Overall, greater recruitment for visual word reading than spontaneous speech**
• **However, higher encoding for complex, spontaneous tasks (sentences) than cued task and overlearned tasks (counting, months)**
• **Spontaneous speech is encoded in signals of IFG and MTG more robustly than cued speech**

References

[1] Bookheimer, S., Zeffiro, T., Blaxton, T., Gaillard, W. & Theodore, W. (2000). Activation of language cortex with automatic speech tasks. *Neurology, 55*(8), 1151-1157.
[2] Brennan, N. M. P. et al. (2007). Object naming is a more sensitive measure of speech localization than number counting: Converging evidence from direct cortical stimulation and fMRI. *NeuroImage, 37*(Suppl. 1), S100–S108.
[3] Hamberger, M.J. (2007). Cortical Language Mapping in Epilepsy: A Critical Review. *Neuropsychol Rev 17*, 477–489.
[4] Flinker, A. et al. (2015). Redefining the role of Broca's area in speech. *Proc. Natl. Acad. Sci, 112*(9), 2871–2875.
[5] Magrassi, L., Aromataris, G., Cabrini, A., Annovazzi-Lodi, V., & Moro, A. (2015). Sound representation in higher language areas during language generation. *Proc. Natl. Acad. Sci, 112*(6), 1868–1873.

Acknowledgments