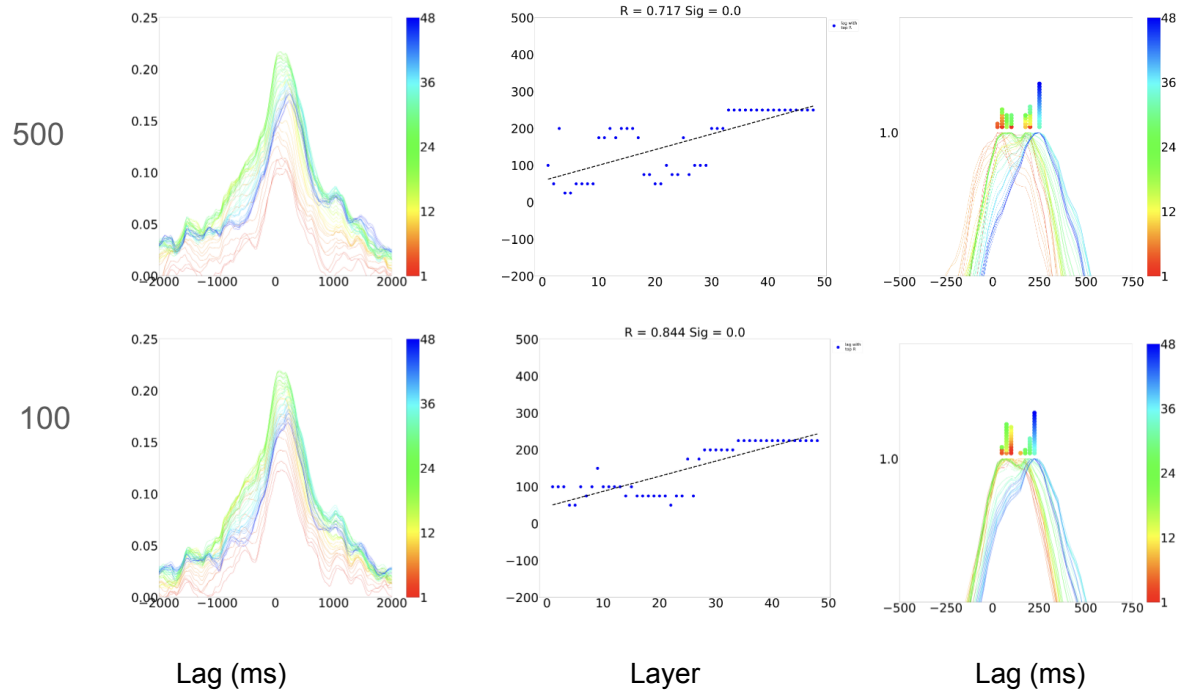


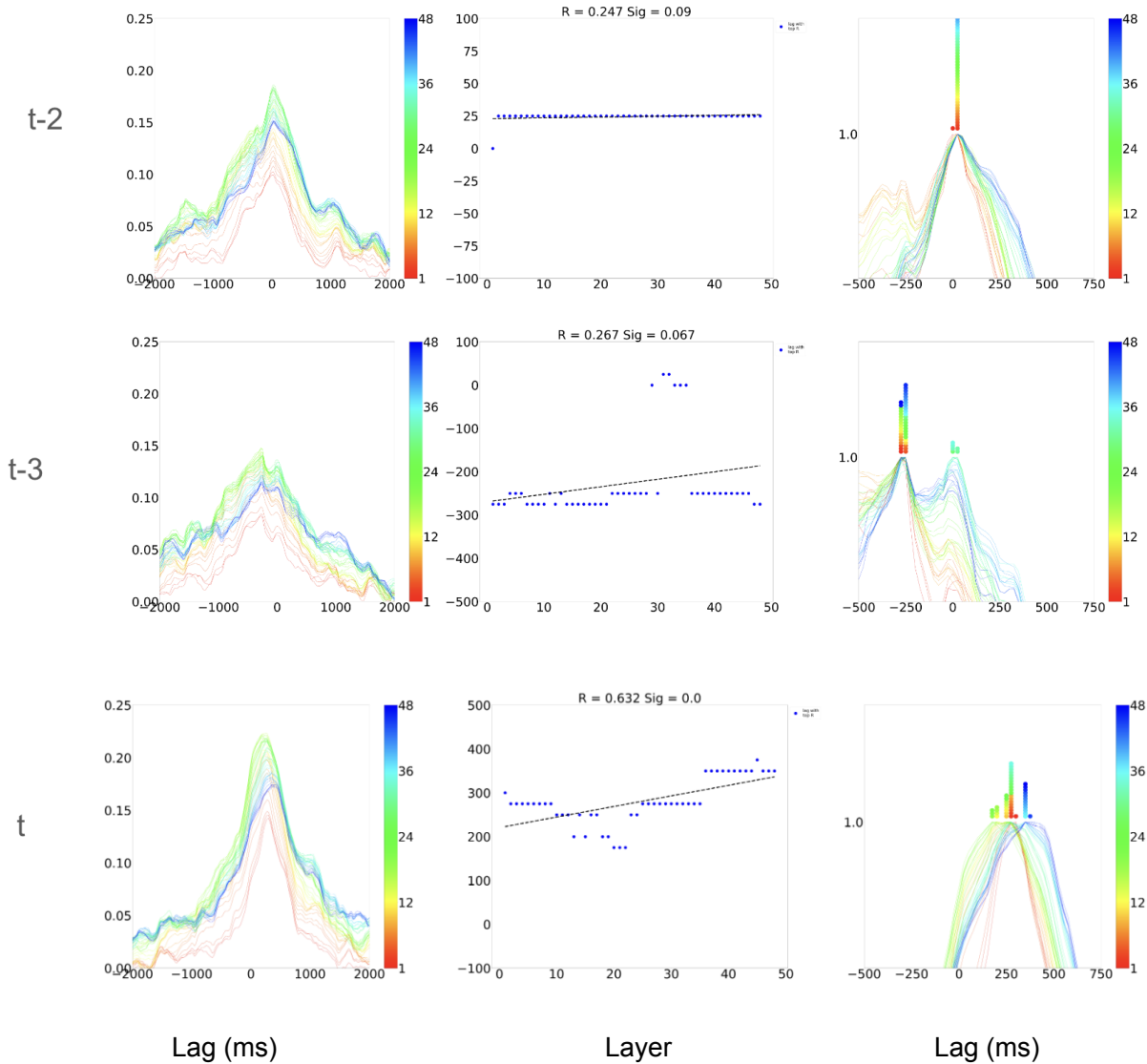
Supplementary Figure S1: Left to right: Scaled encoding, lag-layer and regular encoding plots for Bert large embeddings across ROIs. The y-axis for the regular encoding plot is correlation. The y-axis of the lag layer plot is lag (ms). The words used to train the encoding model were the predictable words.

GPT2-XL With Different Context Lengths

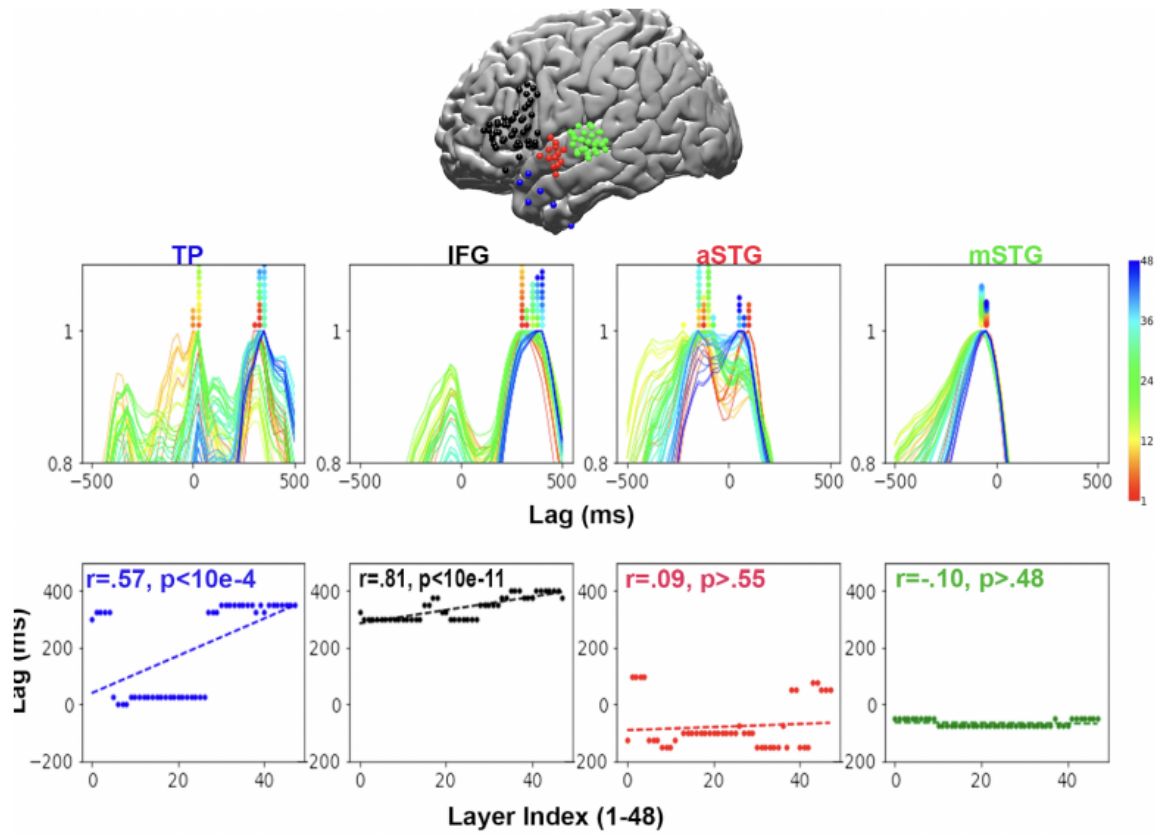


Supplementary Figure S2: Left to right: Encoding, lag layer and scaled encoding plots when GPT2-XL is given input sizes of 500 and 100 words. The y-axis for the regular encoding plot is correlation. The y-axis of the lag layer plot is lag (ms). The ROI observed is the IFG and the encoding model was trained on predictable words.

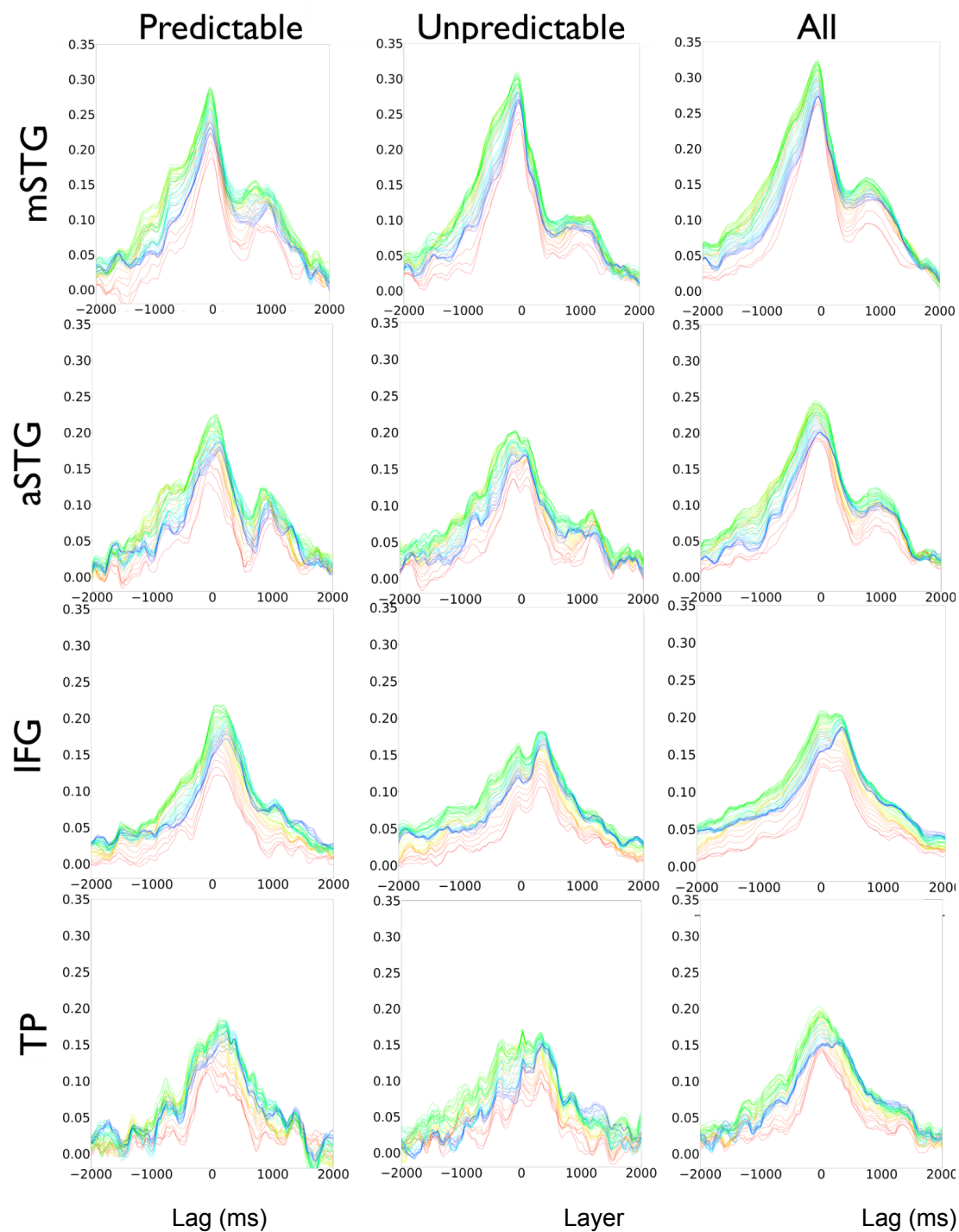
Encoding and Lag Layer Performance Using Embeddings for Adjacent Words



Supplementary Figure S3: Left to right: regular encoding, lag-layer and scaled encoding plots for the IFG and predictable words where the encoding model was given different words for the same neural signal. $T-1$ is what we have in our paper, which is the embedding generated when GPT2-XL is given all words up until the current word. $T-2$ is when the embedding is from the prior word. $T-3$ is the embedding before that. T is when the embedding is the prediction GPT2-XL makes when given the context up until and including the current word. The y-axis for the regular encoding plot is correlation. The y-axis of the lag layer plot is lag (ms).

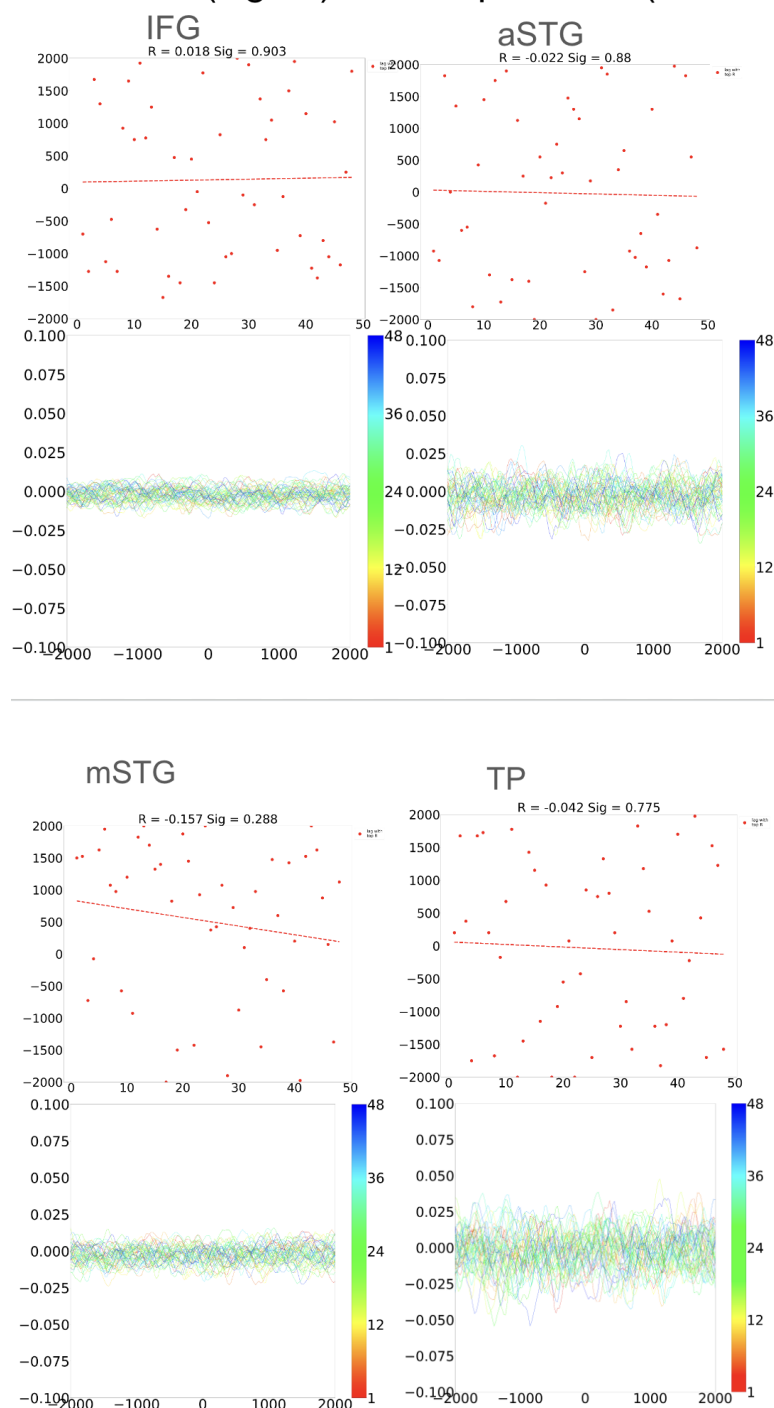


Supplementary Figure S4: Temporal hierarchy along the ventral language stream for unpredictable words. (Top) location of electrodes on the brain, color coded by ROI with blue, black, red, green corresponding to TP, IFG, aSTG and mSTG respectively. (Middle) Scaled encoding performance for these ROIs. (Bottom) Scatter plot of the lag that yields peak encoding performance for each layer.

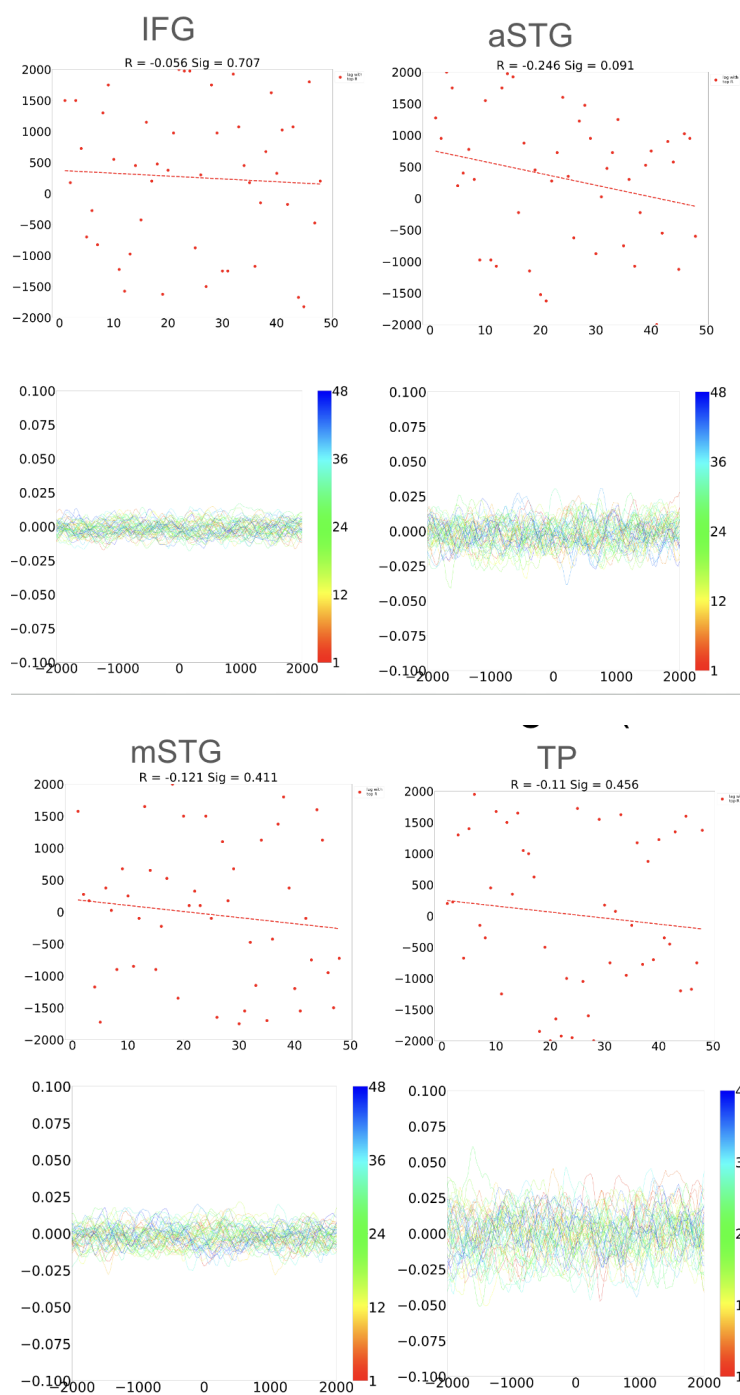


Supplementary Figure S5: Regular encoding plots for combinations of ROIs (mSTG, aSTG, IFG, TP) and word predictability (predictable, unpredictable, all words). The y-axis for the regular encoding plot is correlation.

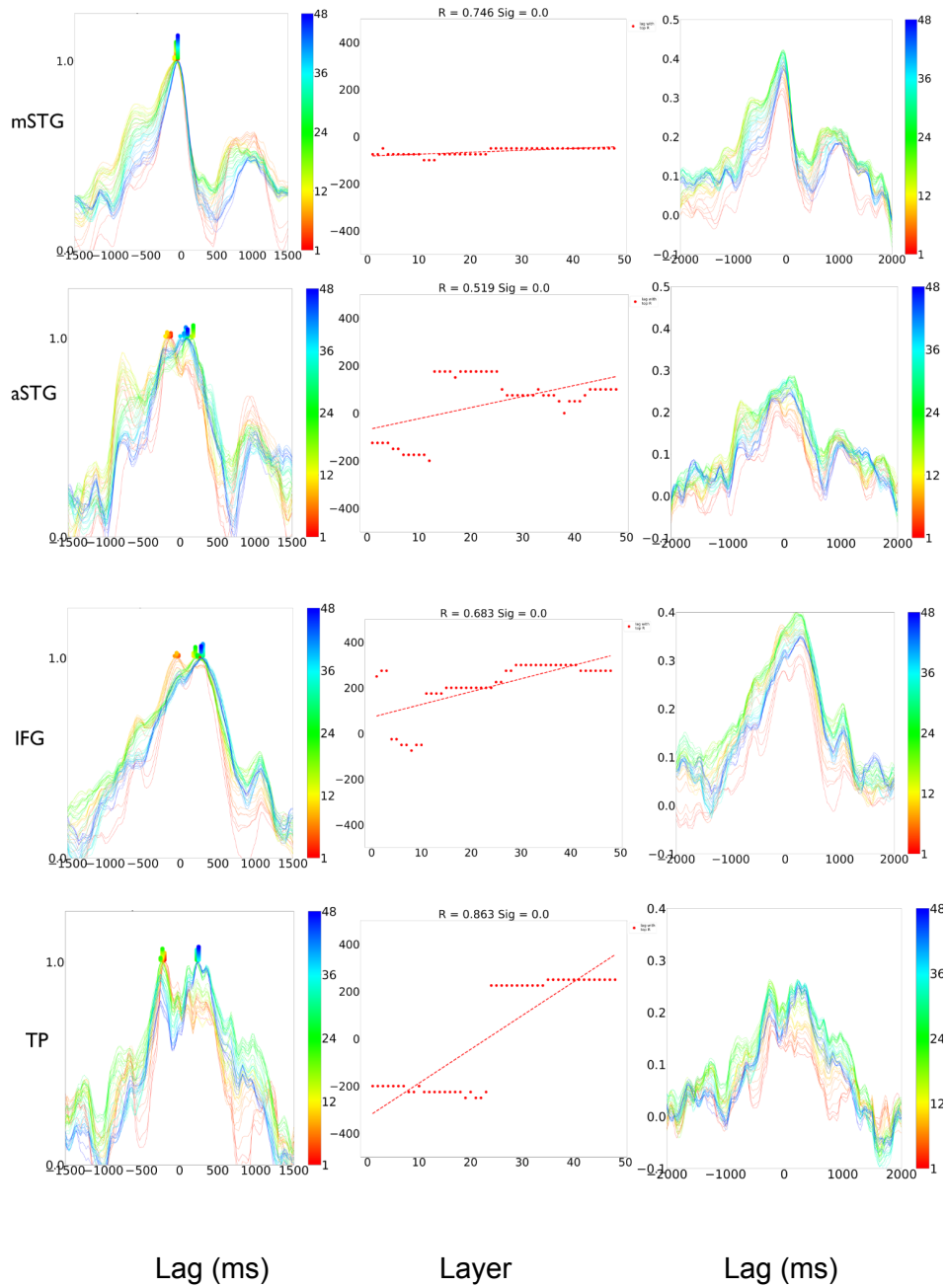
Shuffle Y (signal) with respect to X (embedding)



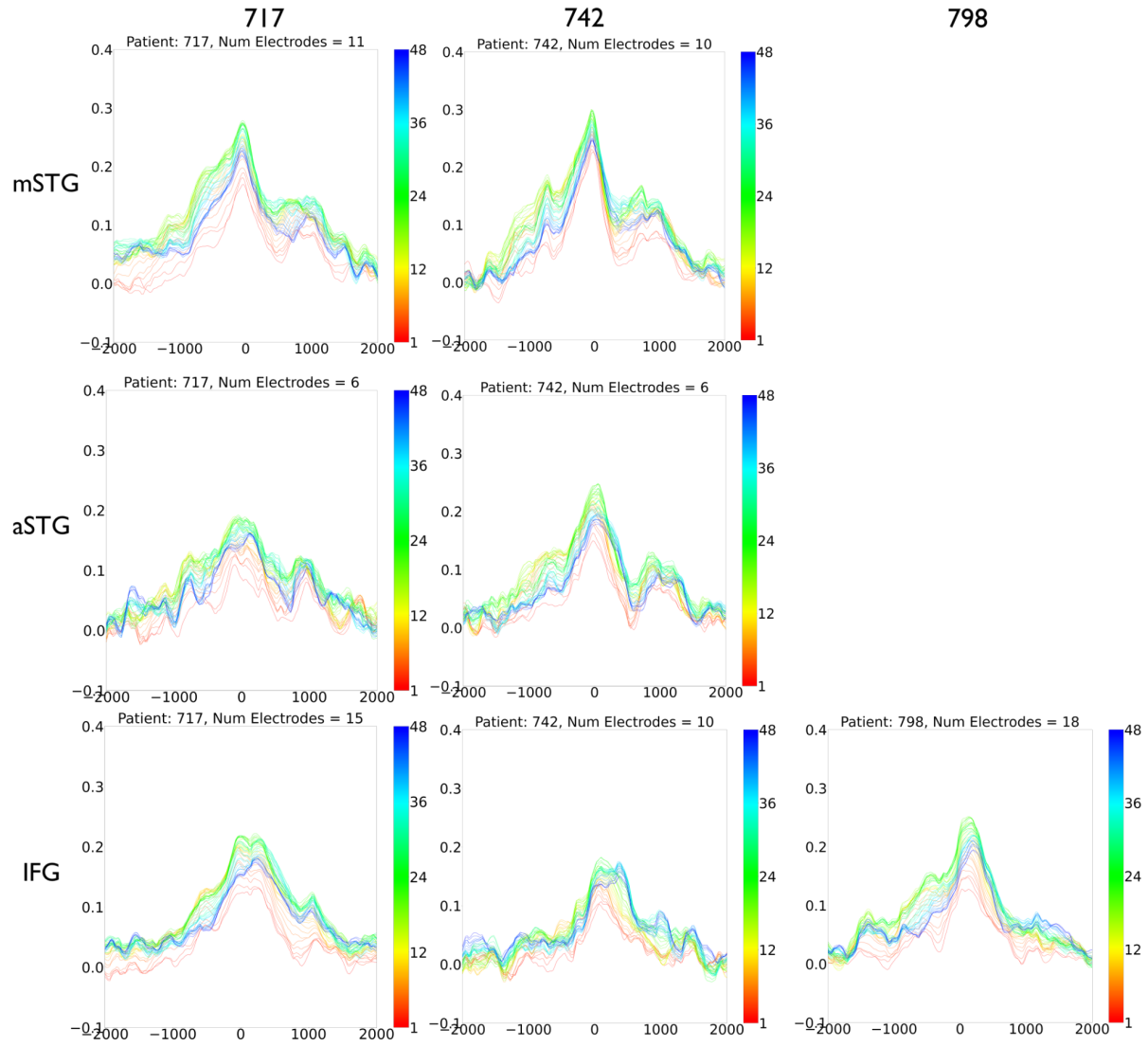
Supplementary Figure S6: Control analysis where, for each ROI we took all encoding training examples $\{(X_i, Y_i)\}$, where $i = 1, \dots$, number of predictable words, X_i : embedding of word i , Y_i : neural signal centered at onset of word i . We shuffled the Y_i 's with respect to the X_i 's. We then re-ran our encoding analyses. The top player shows the result of plotting lag of maximal encoding correlation (y-axis) vs. layer (x-axis) and the bottom shows the encoding correlations for lags between -2000 ms and 2000 ms.



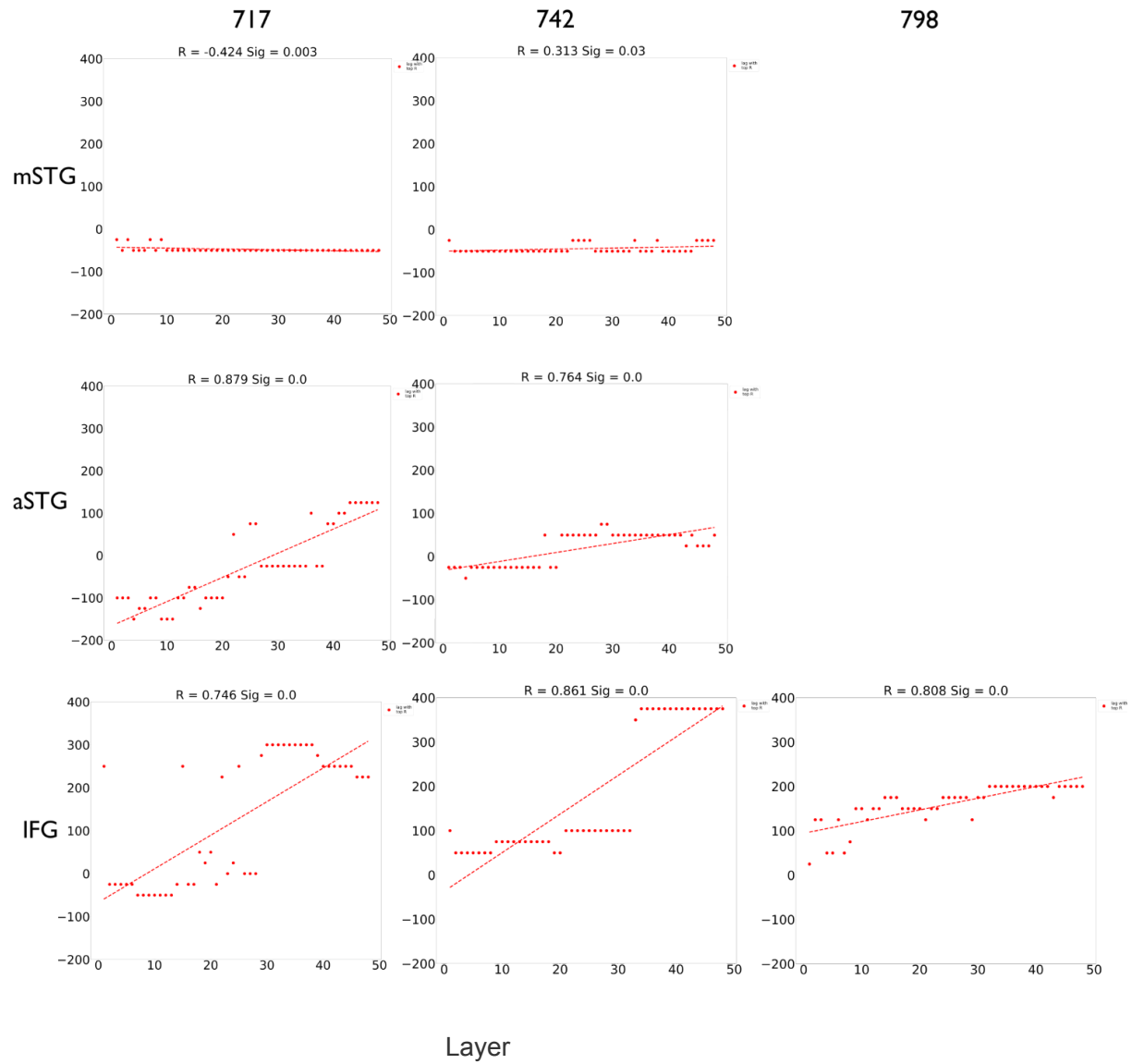
Supplementary Figure S7: Control analysis where, for each ROI we took all encoding training examples $\{(X_i, Y_i)\}$, where $i = 1, \dots$, number of predictable words X_i : embedding of word i , Y_i : neural signal centered at onset of word i . We then phase shuffled the Y_i 's and re-ran our encoding analyses. The top player shows the result of plotting lag of maximal encoding correlation (y-axis) vs. layer (x-axis) and the bottom shows the encoding correlations for lags between -2000 ms and 2000 ms.



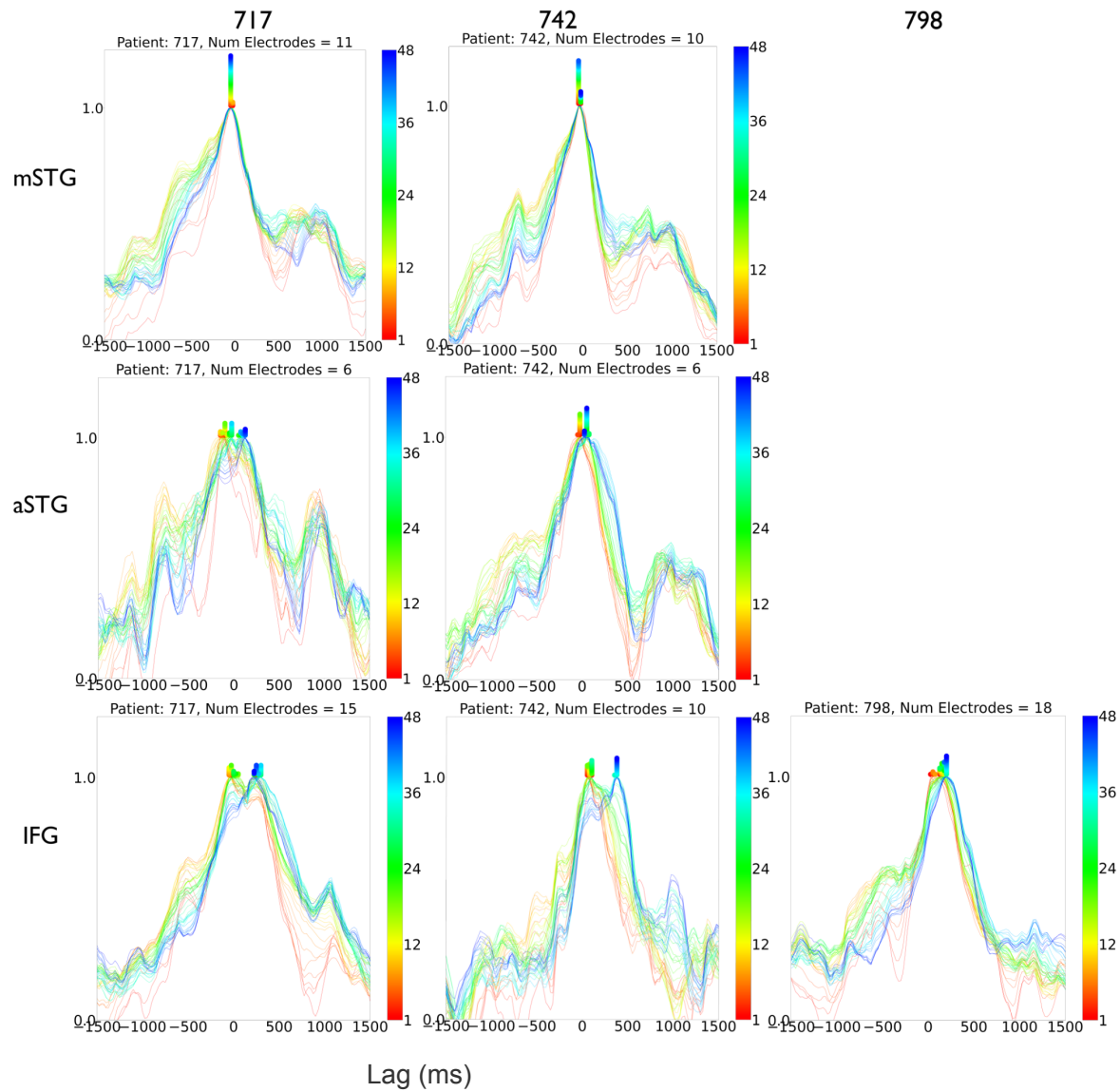
Supplementary Figure S8: Here we re-ran our analyses but trained a single encoding model per ROI, where it was tasked to predict the neural signal averaged over electrodes in that ROI. Left to right: regular encoding, lag-layer and scaled encoding plots for all ROIs, where the encoding model was trained on predictable words. The y-axis for the regular encoding plots is correlation and the y-axis for the lag-layer plots is lag in ms.



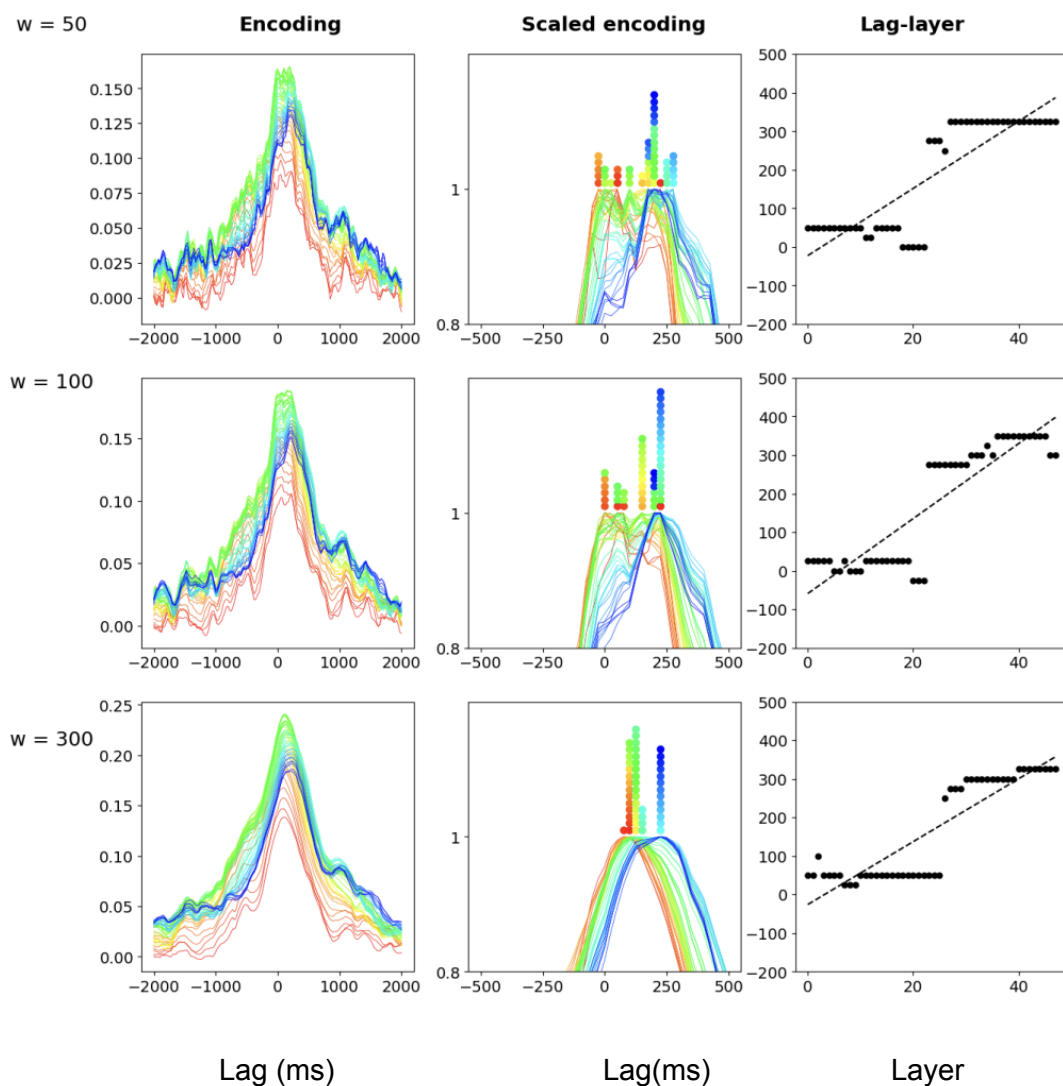
Supplementary Figure S9: Per-patient regular encoding results for patients 717, 742, 798, which were the only patients that had greater than 5 electrodes for any ROIs. No patient had 5 or more significant electrodes in the TP, so we omitted it from this analysis. Patient 798 only had 5 or more (18) electrodes in the IFG. For the numbers of electrodes for each patient-ROI combination, see the subplot titles. Y-axis is correlation.



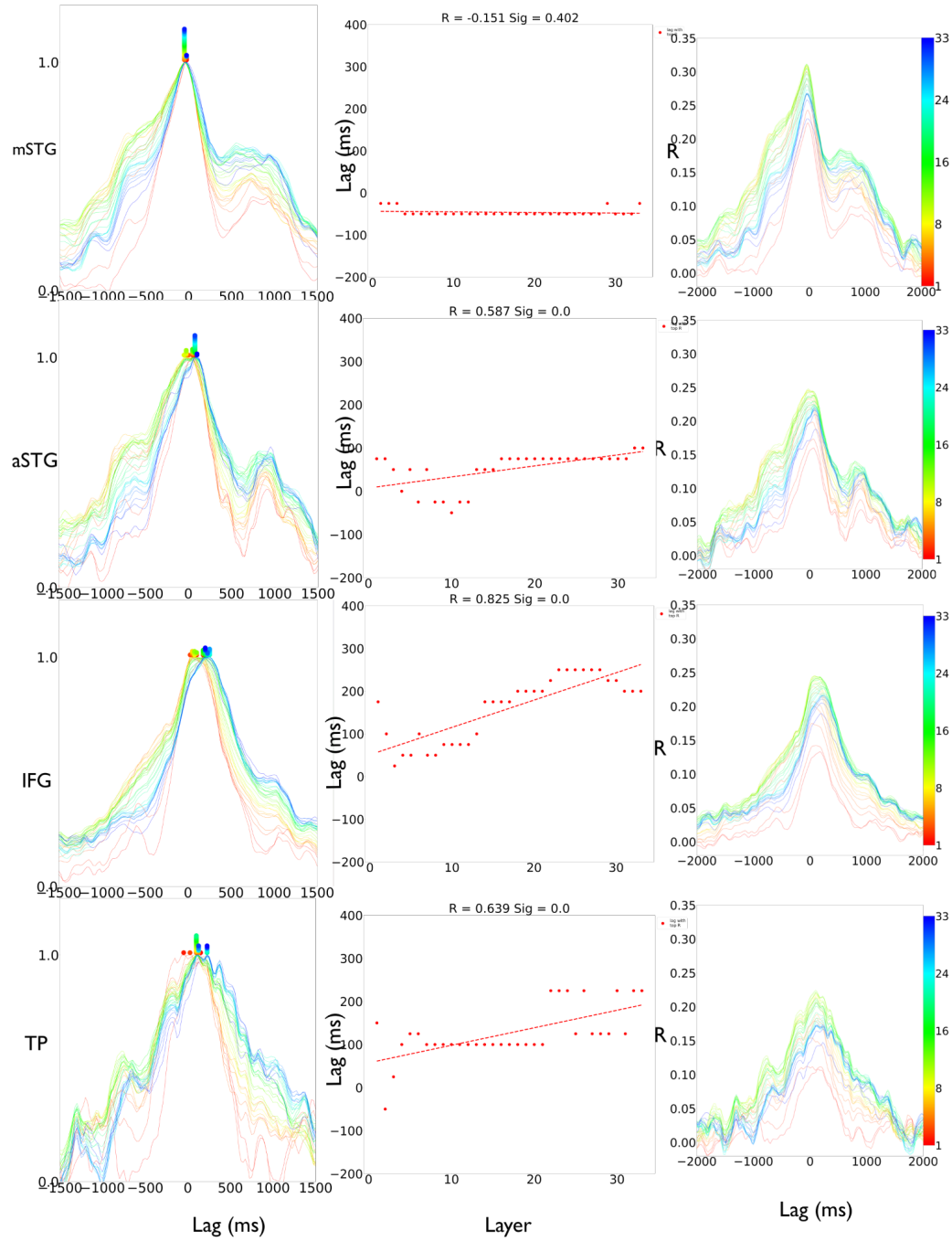
Supplementary Figure S10: Per-patient lag-layer results for patients 717, 742, 798. Y-axis is lag (ms).



Supplementary Figure S11: Per-patient scaled encoding results for patients 717, 742, 798.



Supplementary Figure S12: We re-ran the regular encoding analysis, scaled encoding analysis and lag-layer analysis (left to right) for different smoothing window sizes applied to the neural signal (windows of 50, 100, 300). The correlations are all positive ($r > 0.75$) and significant ($p < 1e-10$). The y axis or regular encoding is correlation. The y-axis of the lag layer analysis is lag (ms).



Supplementary Figure S13: Left to right: Scaled encoding, lag-layer and regular encoding plots for LLaMA 7b embeddings across ROIs. The words used to train the encoding model were the predictable words for GPT2-XL.

Table 4: Patient demographic information

Patient ID	Age	Gender	Ethnicity	Language
798	24	M	White	English
763	42	M	Other	English primary; Spanish secondary
743	23	F	White	English
742	34	F	Hispanic	English/Spanish
741	24	M	White	English
723	25	M	White	English
717	48	F	Irish	English

Supplementary Table T1: Patient demographic information.

Table 1: Additional information about patient pathology and neuropsychological scores.

NO.	Neuropsych Score	Pathology/Epilepsy type/Seizure Focus	Implant
798	VCI: 145 POI: 96 PSI: 86 WMI: 95	Focal epilepsy arising from the left hemisphere with a broad focus involving the left temporal neocortex (superior, middle, inferior temporal gyri) left frontal operculum, inferior postcentral gyrus, insula	Left grid and strips
763	VCI: 87 POI: 123 PSI: 97 WMI: 89	Left temporal lobe epilepsy. Ictal onsets localized to the left temporal lobe (perilesionally) and left posterior mesial temporal lobe.	Left grid, strips, and depth electrodes
743	VCI: 100 POI: 109 PSI: 92 WMI: 86	Focal epilepsy localized to the left posterior insula and periopercular region at the frontoparietal junction.	Left grid, strips, and depth electrodes
742	not found	Probable focal epilepsy, not clearly lateralized or localized. ICEEG showed definitively that disabling clinical events were psychogenic non-epileptic attacks.	Left grid, strips, and depth electrodes
741	VCI: 96 POI: 79 PSI: 81 WMI: 86	Left hemispheric multilobar epilepsy	Left grid, strips, and depth electrodes
723	not found	Bilateral mesial temporal lobe epilepsy	Bilateral strips and depth electrodes
717	VCI: 107 POI: 104 PSI: 111 WMI: 114	Right anteromesial temporal lobe epilepsy. ICEEG localized ictal onsets to the right temporal pole and right hippocampus	Bilateral strips and depths, and a left grid

Supplementary Table T2: Patient clinical information.