Supplementary Information for
Frontal cortex organization supporting audiovisual processing during naturalistic viewing

Faxin Zhou, Amirhossein Khalilian-Gourtani, Patricia Dugan, Andrew Michalak, Orrin Devinsky, Peter Rozman, Werner Doyle, Daniel Friedman, Adeen Flinker

Faxin Zhou: fz2185@nyu.edu; Adeen Flinker: adeen@nyu.edu

Supplementary Information

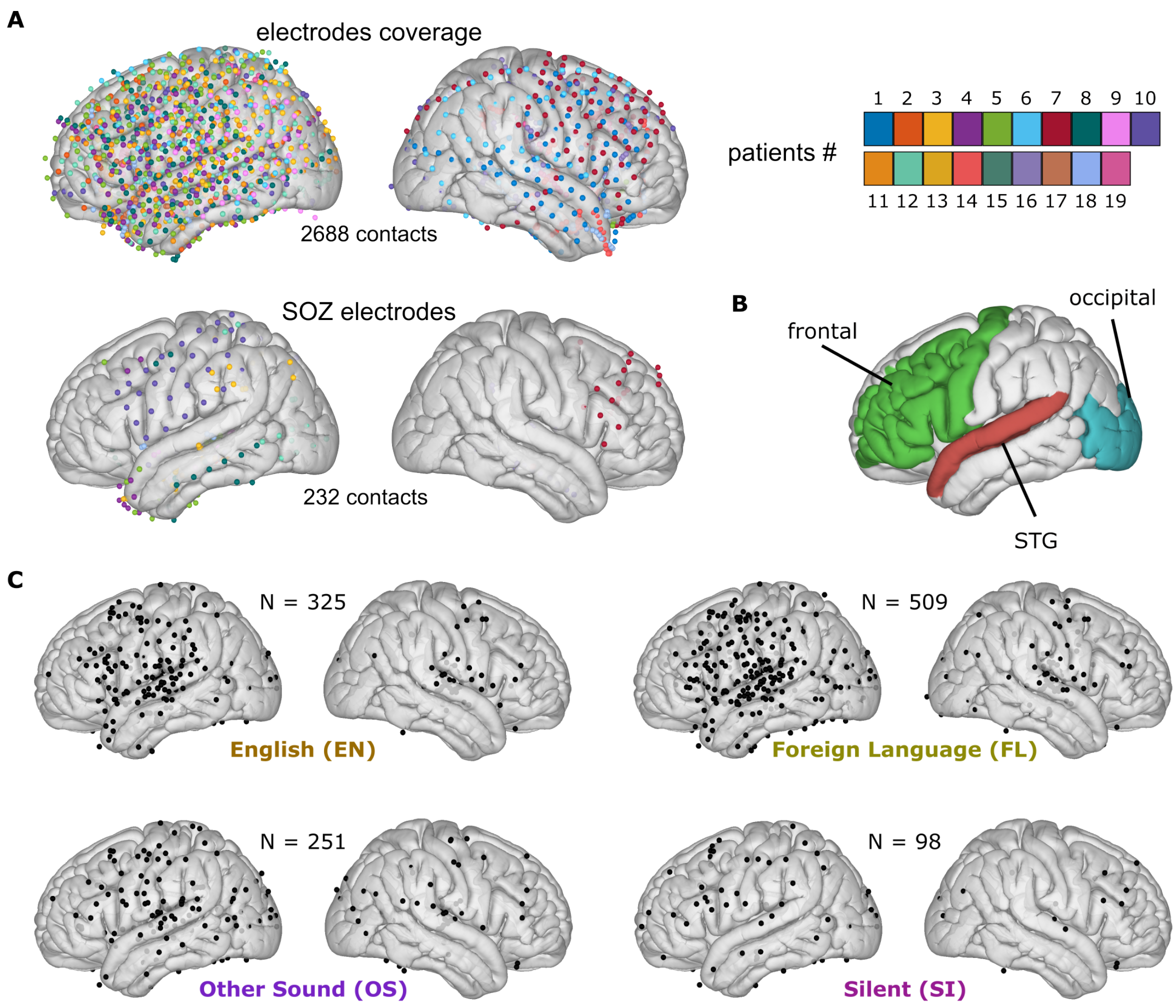

Figure S1: (A) The electrode coverage (upper panel; 2688 contacts in total) and the seizure onset zone (SOZ) electrode distribution (lower panel; 232 contacts in total which were removed from analysis including depth and surface) for all 19 participants involved in the movie viewing task. (B) Three ROIs used in this study, including the superior temporal gyrus (STG), frontal regions (comprising the inferior frontal gyrus, middle frontal gyrus, and pre-central gyrus), and the occipital regions (comprising the lateral occipital gyri, lingual gyrus, and cuneus gyrus). (C) The brain maps of the active electrodes for all four conditions that survived the permutation test (Fig. S2; see Methods: Active electrodes selection).

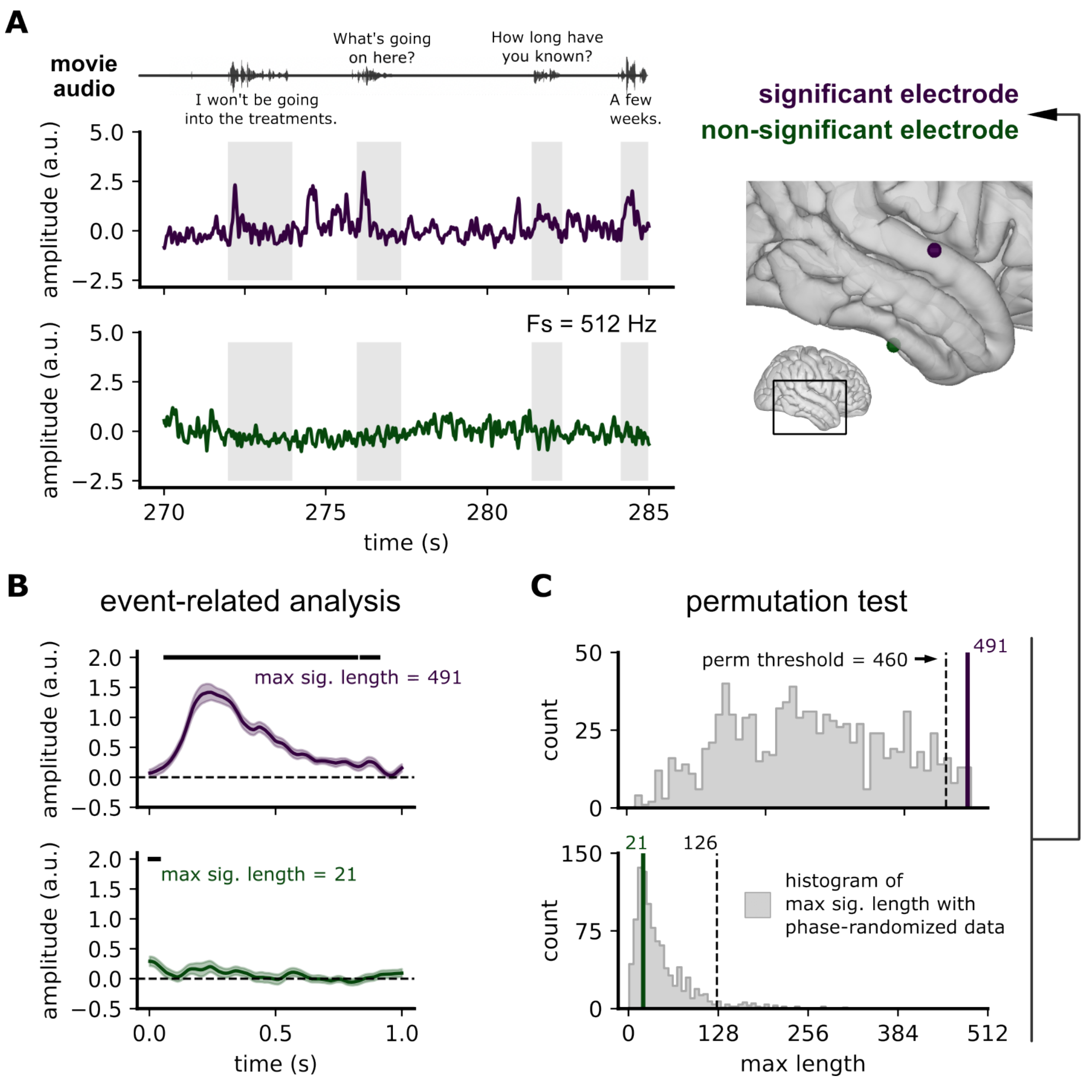

Figure S2: **Permutation procedure for selecting the active electrodes.** (A) Representative signals from two electrodes (right panel): one located in the STG (purple) and the other located in the ITG (green). The STG electrode exhibited responses to auditory events (left panel), whereas the ITG electrode showed no response. (B) Event-related analysis. The neural activity during the events was aligned to the onset and averaged across time (the temporal window was set as 1 seconds). Then, the statistical significance was determined for each time point using a two-sided t-test with $p < 0.05$. The maximum length of the consecutive significant time points (i.e., maximum significant length (L_{max})) is recorded. (C) Permutation test. We repeated the event-related analysis on the phase-randomized signals for 1000 times, and computed the maximum significant length for each permutation (L^p_{max}). This procedure generated a null distribution of the L^p_{max}. The real length L_{max} was then compared against this null distribution, with significance determined at the top 5th percentile threshold (FDR corrected).

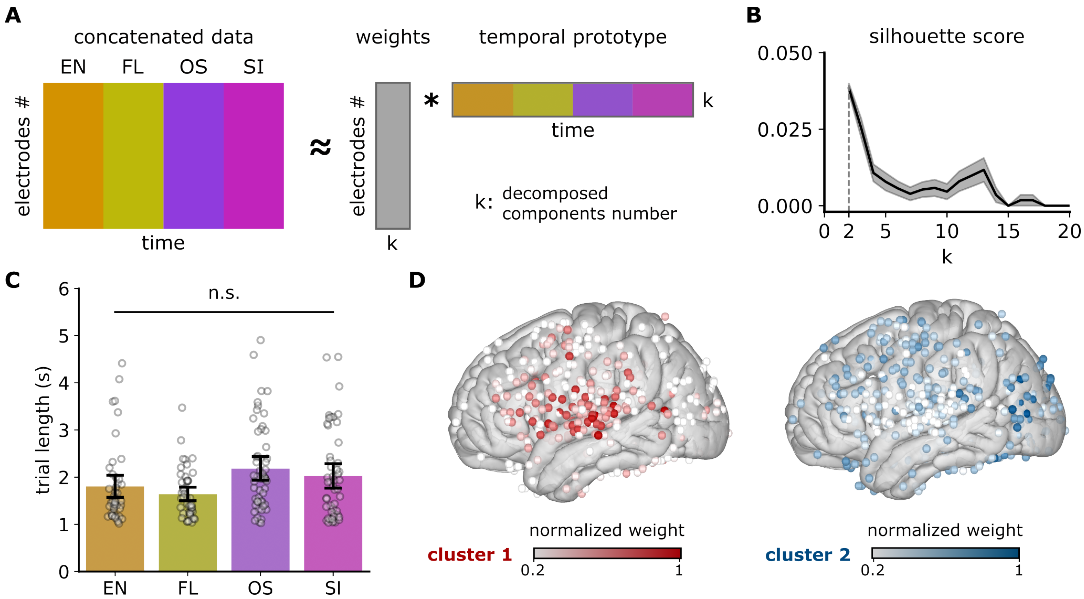

Figure S3: **Unsupervised functional clustering.** (A)The data for all four conditions (English (EN), Foreign Language (FL), Other Sound (OS), and Silent (SI)) were concatenated and then decomposed into a weight matrix and a temporal prototype matrix (see Methods: Functional clustering analysis). (B) The result of silhouette method used to determine the optimal number of clusters (k). Higher scores indicate better clustering performance. The shaded area around the curve represents the standard error (SE). (C) Bar plots of trial lengths across four conditions (EN: $n = 44$ trials; FL: $n = 46$ trials; OS: $n = 58$ trials; SI: $n = 51$ trials). Each dot represents an individual trial. (D) The brain maps for the two NMF clusters separately.

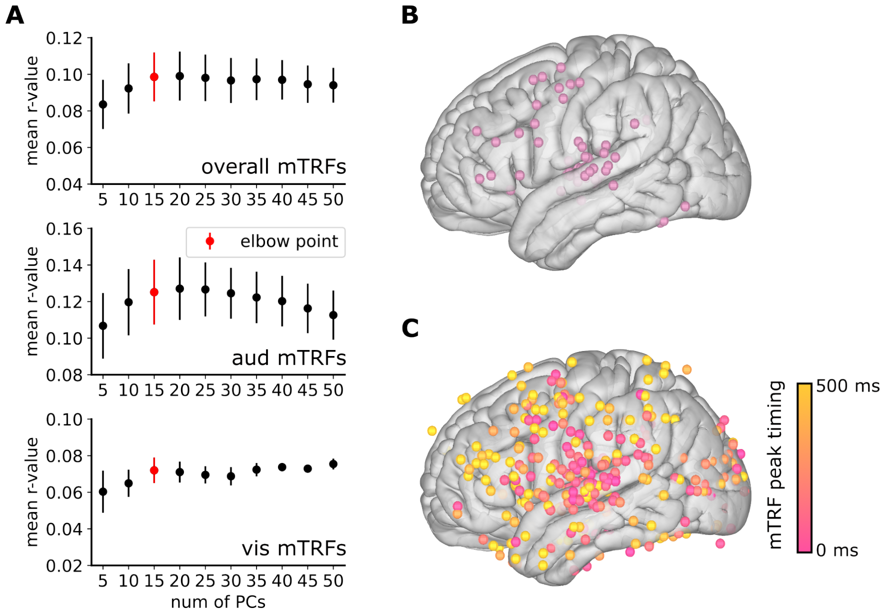

Figure S4: (A) Results of the incremental principal components (PCs) test across all the mTRF models (upper panel), auditory mTRF models (middle panel), and visual mTRF models (lower panel). The red error bars indicate the elbow points identified using kneed Python toolbox, which employs a rotation-based algorithm to detect the maximum curvature [1]. (B) Brain map showing significant electrodes for both audio and visual mTRF models. (C) Timing of peak weights across electrodes that showed significant responses in the auditory and visual mTRF models.

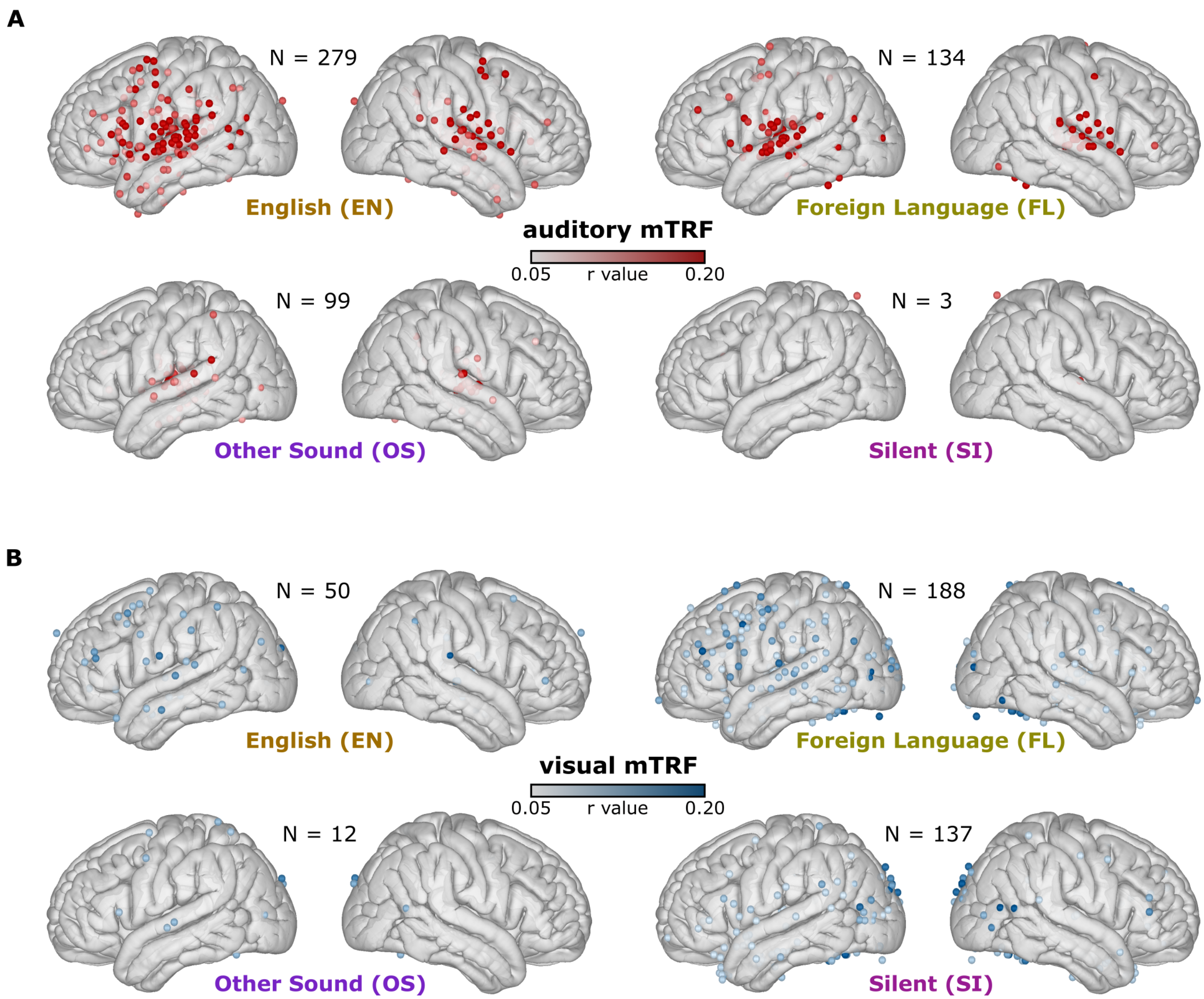

Figure S5: Results of encoding models for auditory (A) and visual (B) information in four conditions. N represents the number of significant electrodes.

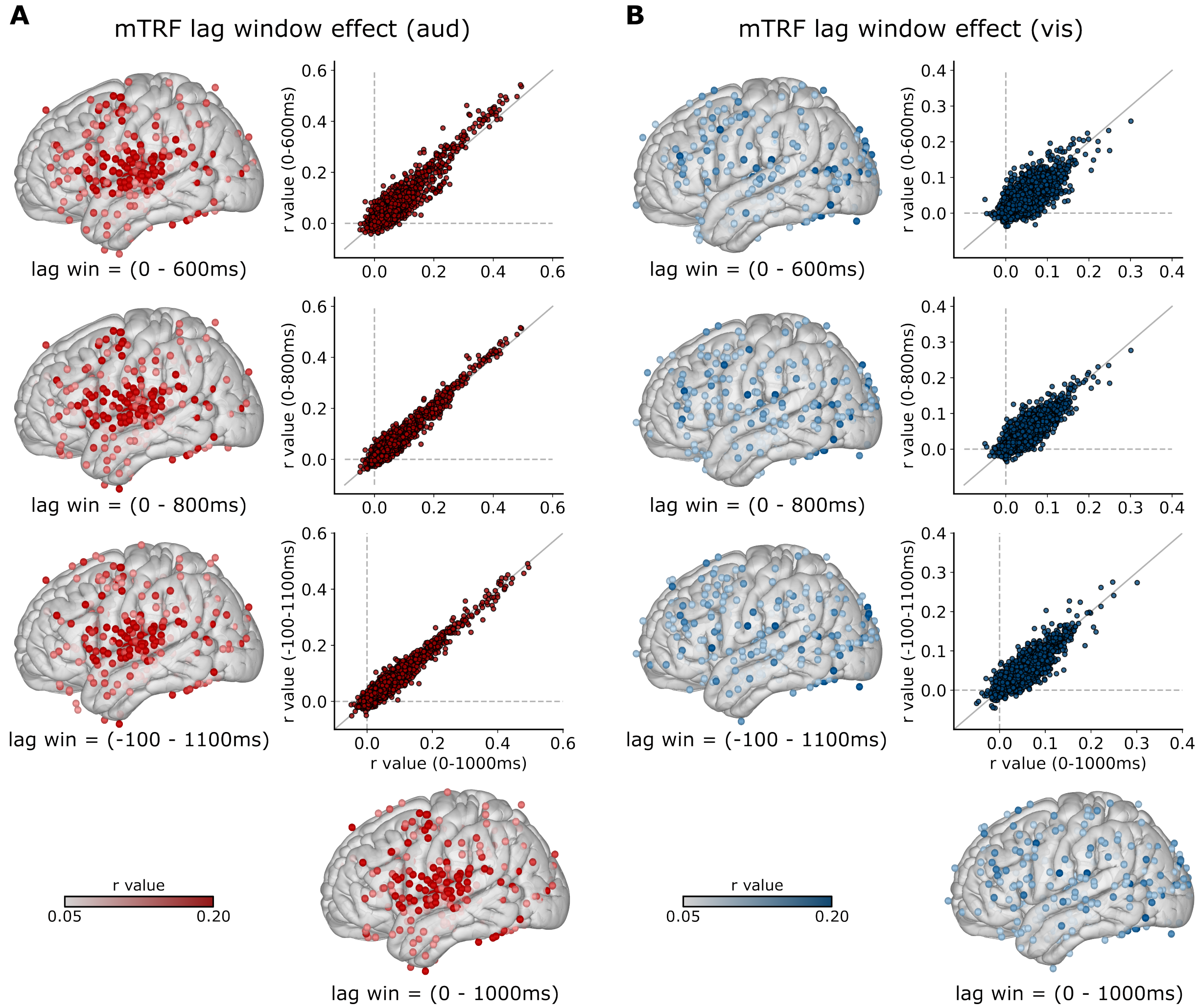

Figure S6: Brain maps and r-value comparisons across increasing lag windows (0–600 ms, 0–800 ms, 0–1000 ms, and –100–1100 ms) for both auditory (A) and visual (B) mTRF models. No permutation test or multiple correction were applied; a hard threshold of $r > 0.1$ was used for visualization.

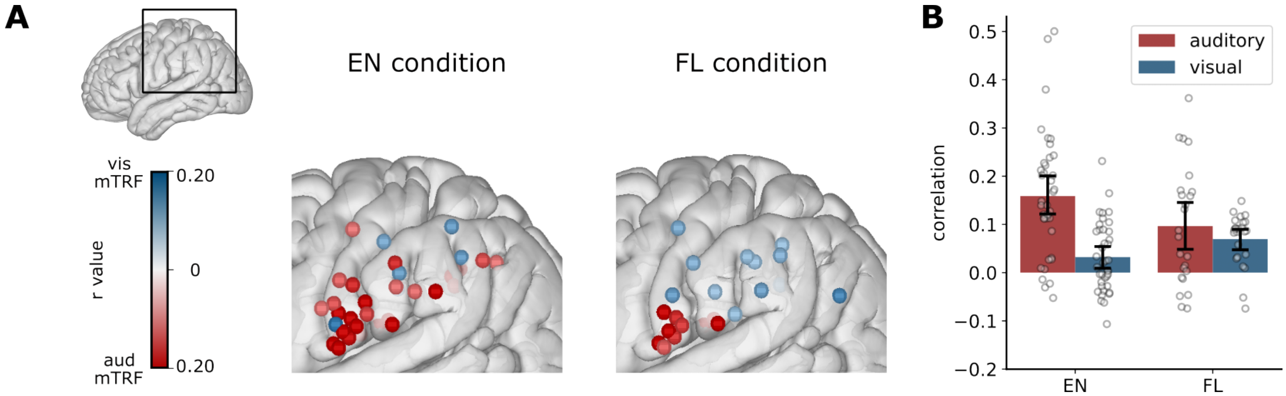

Figure S7: (A) The mTRF correlation maps in the parietal regions (including the postcentral gyrus, inferior parietal lobe, and supramarginal gyrus) for the English (EN) and Foreign Language (FL) conditions. (B) Bar plots of the auditory and visual mTRF models in the parietal cortex for the EN and FL conditions. Error bars represent 95% confidence interval (CI).

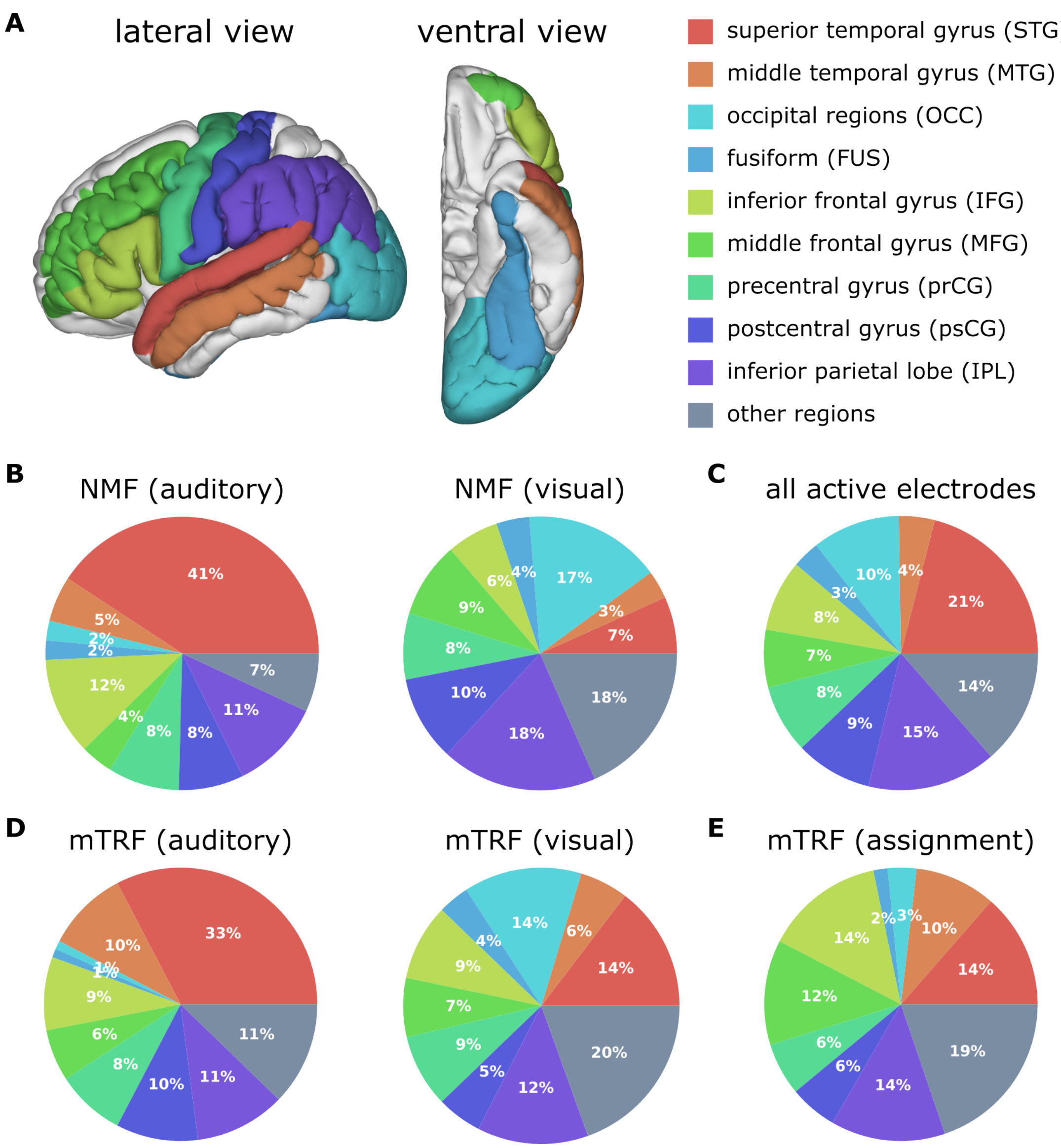

Figure S8: **Anatomical distributions for the NMF, mTRF and active electrodes.** (A) Nine ROIs illustrated on the brain. (B) Pie charts showing the anatomical distribution of electrodes from the auditory and visual functional clusters (NMF approach). (C) Pie charts showing the anatomical distribution of the active electrodes. (D) Pie charts depicting the anatomical distribution of significant electrodes from the auditory and visual encoding models (mTRF approach). (E) Pie chart showing the anatomical distribution of significant electrodes responsive to the audiovisual assignment (i.e., the global modality). The total number of electrodes within each ROI, and the number (and percentage) of electrodes for figure B-E are shown in Table S2.

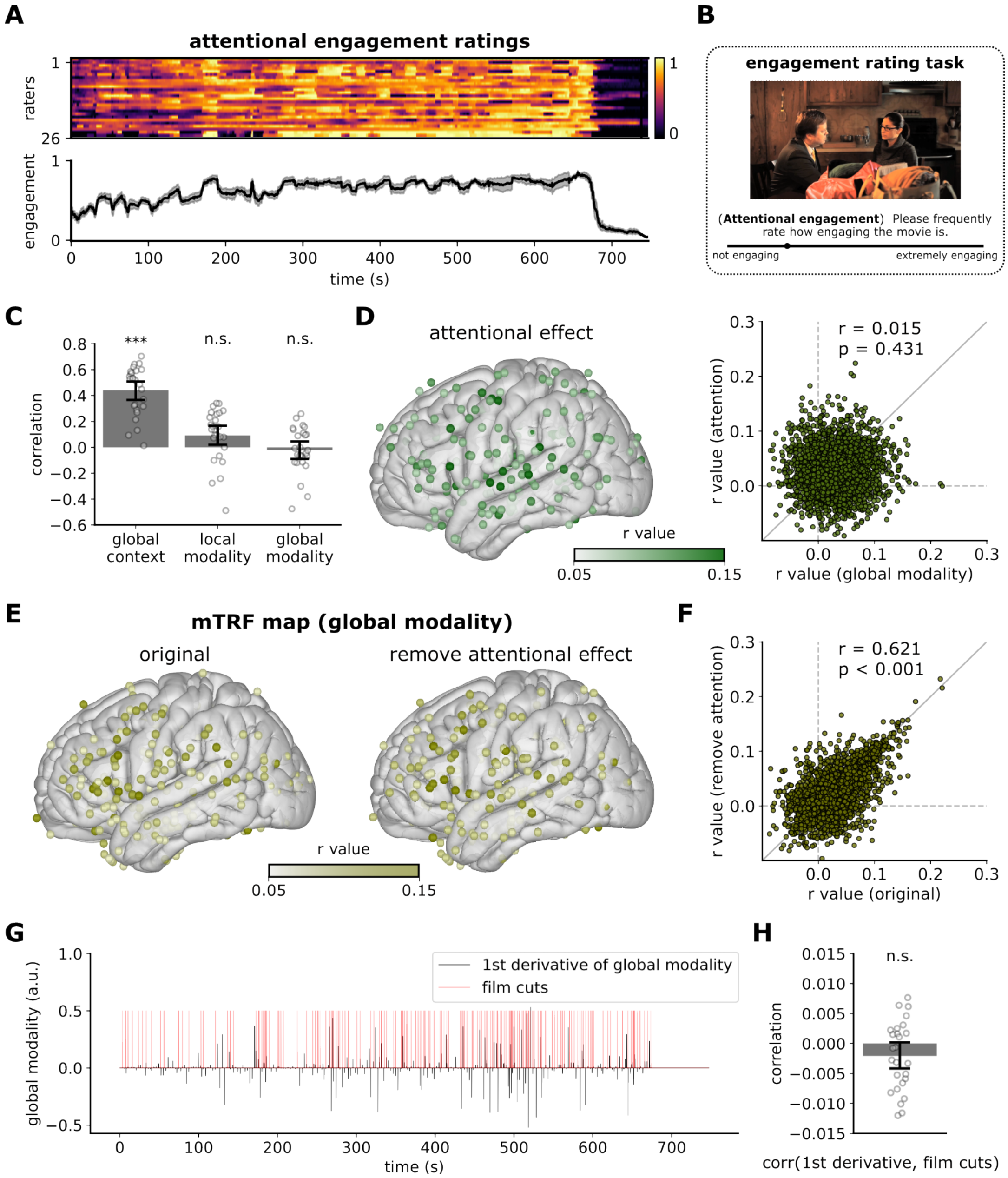

Figure S9: (A) The attentional engagement ratings from all participants over time (upper panel) and the average time series across participants (lower panel). (B) Illustration of the engagement rating task conducted on the Amazon Mechanical Turk platform. (C) Correlations between attentional engagement ratings and global context (one-sample t test; $T(25) = 11.987$, $p < 0.001$, $d = 2.351$), local modality (one-sample t test; $T(25) = 2.068$, $p = 0.074$, $d = 0.406$), as well as global modality (one-sample t test; $T(25) = -0.608$, $p = 0.549$, $d = 0.119$). Each dot represents an individual rater. (D) The mTRF result for the attentional engagement effect. (E) The mTRF results for the original global modality (left) and after removing the attentional effect (right). (F) Comparison of r values between the two models across all electrodes. (G) Markers of abrupt changes in the averaged global modality (black; quantified as its first derivative) and the film cuts (red). (H) Correlations between the film cuts and the global modality changes across participants. Each dot represents an individual rater. Error bars in figure C and H represent 95% CI.

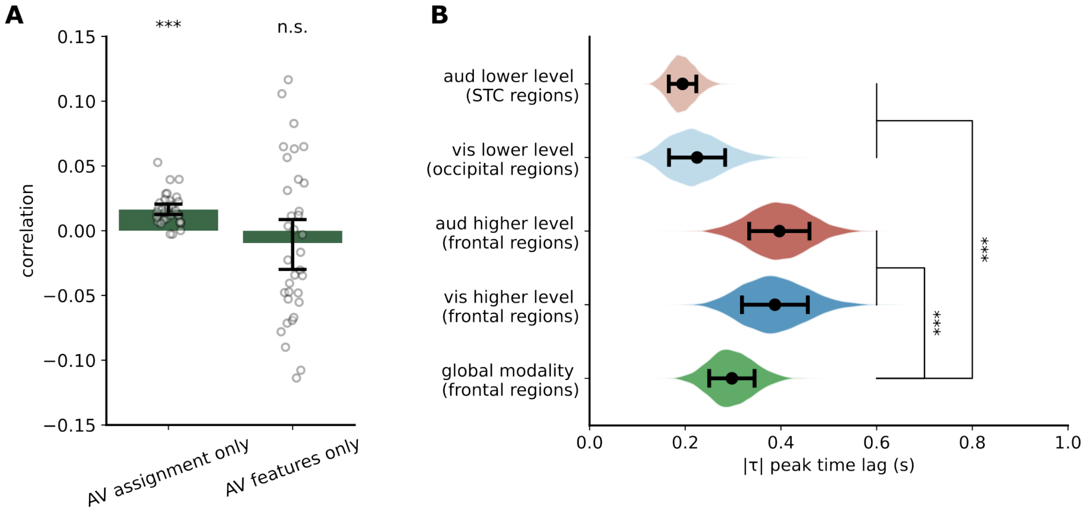

Figure S10: (A) The unique contributions for modality assignment and audiovisual features in frontal electrodes shown in Fig. 6A (see Methods: Encoding model partitioning). A significant effect was observed only for the audiovisual assignment (one sample t-test, $t(33) = 7.968$, $p < 0.001$, $d = 1.366$, FDR corrected) but not for the audiovisual features (one sample t-test, $t(33) = -0.927$, $p = 0.361$, $d = 0.159$, FDR corrected). The error bars represent 95% CI. (B) Timing analysis of audiovisual (AV) assignment. The timing of AV assignment in the frontal cortex (297 ± 47 ms) was significantly later than the neural representation of AV features in the perceptual regions. A significant difference was also observed between the timing of encoding the AV features and the modality assignment in the frontal regions. The distributions were estimated based on a bootstrap procedure, and the statistical significance between regions was assessed via a permutation test (see Methods: Timing analysis of encoding models). The error bars represent the standard deviation (SD). Significance levels are set as $p < 0.001$ (***), $p < 0.01$ (**), $p < 0.05$ (*), and $p \geq 0.05$ (n.s.).

Table S1: Patient information

ID	Coverage	Gender	Handedness	Age at Surgery (y)	Seizure Onset Zone (SOZ)
1	Right	Male	Right	29	temporal
2	Left	Female	Right	45	insula/frontal
3	Left	Male	Right	24	parietal
4	Left	Female	Right	30	temporal
5	Left	Female	Right	56	temporal
6	Bilateral	Male	Right	26	temporal
7	Right	Female	Right	21	frontal
8	Left	Male	Right	22	temporal
9	Left	Male	Right	44	temporal
10	Left	Female	Right	33	insula/temporal
11	Left	Female	Right	23	temporal
12	Left	Female	Right	18	occipital
13	Left	Female	Right	61	temporal
14	Bilateral	Female	Right	33	temporal
15	Left	Female	Right	27	temporal/frontal
16	Right	Male	Right	28	insula/temporal/frontal
17	Left	Male	Right	26	parietal
18	Bilateral	Female	Right	33	insula/temporal
19	Bilateral	Female	Right	30	insula/temporal/frontal

Table S2: Electrode anatomical coverage and significant responses

ROI	total	active	NMF (aud)	NMF (vis)	mTRF (aud)	mTRF (vis)	mTRF (assign.)
STG	146	72 (49.32%)	53 (36.39%)	12 (8.22%)	64 (43.84%)	36 (24.66%)	24 (16.44%)
MTG	136	20 (14.71%)	7 (5.15%)	6 (4.41%)	19 (13.97%)	14 (10.29%)	17 (12.50%)
OCC	59	32 (54.24%)	3 (5.08%)	29 (49.15%)	2 (3.39%)	34 (57.63%)	6 (10.17%)
FUS	22	10 (45.45%)	3(13.64%)	7 (31.82%)	2 (9.09%)	9 (40.91%)	3 (13.64%)
IFG	90	29 (32.22%)	15 (16.67%)	11 (12.22%)	17 (18.89%)	22 (24.44%)	25 (27.78%)
MFG	145	24 (16.55%)	5 (3.45%)	16 (11.03%)	12 (8.28%)	17 (11.72%)	22 (15.17%)
prCG	101	28 (27.72%)	11 (10.89%)	14 (13.86%)	16 (15.84%)	21 (20.79%)	11 (10.89%)
psCG	106	36 (33.96%)	10 (9.43%)	18 (16.98%)	19 (17.92%)	13 (12.26%)	10 (9.43%)
IPL	208	51 (24.52%)	14 (6.73%)	33 (15.87%)	21 (10.10%)	32 (15.38%)	24 (11.54%)
Other regions	293	49 (16.72%)	9 (3.07%)	33 (11.26%)	24 (8.19%)	48 (16.38%)	35 (11.95%)
Other (WM & UNK)	1150	299 (26%)	35 (3.04%)	19 (1.65%)	123 (10.70%)	44 (3.83%)	129 (11.22%)

"ROI" corresponds to the anatomical regions shown in Fig. S8, and "Other (WM & UNK)" indicates other electrodes labeled as white matter (WM) and unknown (UNK). The column "active" reports the number (and percentage) of active electrodes that survived the permutation test (see Fig. 1B). Columns "NMF (aud)" and "NMF (vis)" report the number (and percentage) of electrodes assigned to the auditory and visual functional clusters, respectively. Columns "mTRF (aud)", "mTRF (vis)", and "mTRF (assign.)" report the number (and percentage) of electrodes showing significant encoding for the auditory, visual, and audiovisual assignment models, respectively.

Supplementary References

[1] Ville Satopaa, Jeannie Albrecht, David Irwin, and Barath Raghavan. Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior. In 2011 31st International Conference on Distributed Computing Systems Workshops, pages 166–171, Minneapolis, MN, USA, June 2011. IEEE.