Brain network representation of the topic in speech processing

Adam Boncz, Brigitta Tóth and István Winkler

Research Centre for Natural Sciences, Budapest

Questions / comments are welcome at boncz.adam@ttk.hu

ttk
INSTITUTE OF COGNITIVE NEUROSCIENCE AND PSYCHOLOGY

Background

Speech is usually ambiguous, its meaning only specified by the whole context (*1*). In terms of a coherent speech segment, context can be subdivided into two classes. One is the immediate context composed of the speech segment itself up to the current point of time - we refer to this type of context as the *story* (also called the narrative, *2-3*). The story is formed by integrating the content of speech over time, chunking memory traces into a chain of events (*4*). The other type of context consists of long-term knowledge evoked by the overall situation and the story so far. It consists of schemas, personal experiences, and general world knowledge (*5-6*). We refer to this latter type of context as the *topic*.
Here we focus on how the brain treats *topic* by studying the brain networks differentially sensitive to the topic of relatively long coherent speech segments.

We hypothesize that 1) similar large-scale networks are activated throughout a single-topic speech segment across both listeners and time within the segment and 2) different *topics* activate topologically/topographically different networks, consistently within each listener.

Methods

Task: Participants listened to four similar speech recordings of descriptive newspaper articles on different topics.

- Data recorded for a recently published study (*9*), $N = 26$
- Spech recordings were matched in terms of length (ca. 6 mins), speaker delivery, quality, position within the experimental session, task instructions, and affective valence
- EEG was recorded (64 channels) – standard preprocessing, followed by source reconstruction with anatomical template, 3-layer BEM, and sLORETA, then averaged in 62 cortical areas
- Functional connectivity (FC) was estimated with ciPLV (10) across all areas in 10-s long, non-overlapping segments (40 epochs/topic)
- FC map of each participant and topic was thresholded with a surrogate data-based approach to further reduce the effect of spurious connections (see Figure 1)

General topic sensitivity was tested by estimating with Pearson's correlation the similarity between FC matrices (pairing cells of identical position from the two matrices) across all epochs, both within and across the different speech recordings. This analysis was performed both for the group averages (shown in Figs. 2A and B) and, separately, for each participant.

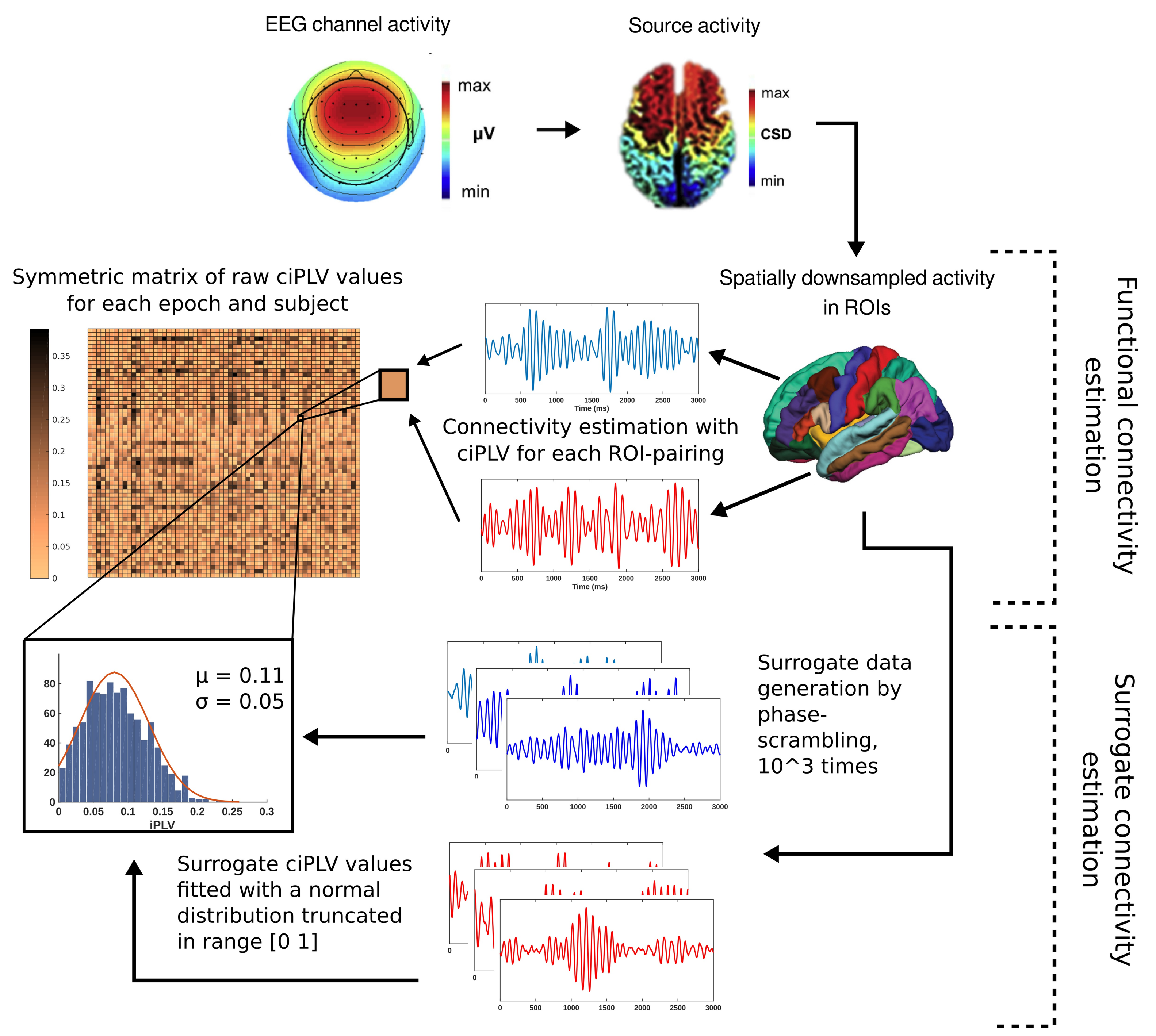

Fig. 1. Overview of the FC estimation pipeline. Source-reconstructed signals were first pooled into 62 cortical regions-of-interest (ROIs). Next, for each epoch of each participant, FC for all ROI-pairings was calculated as the corrected imaginary phase locking value (ciPLV), a measure insensitive to volume conduction (zero lag synchronization). For each ROI-pairing, the distribution of surrogate ciPLV values was estimated from 1000 phase-randomized signals. Actual ciPLV values were compared to the surrogate distributions across all epochs with a random permutation test, then corrected with False Discovery Rate (FDR, $q = .05$).

Results

We show here results for the alpha (8-12 Hz) EEG band. The other bands showed similar results.

1) Group average FC matrices for epoch-pairings within topics were more similar to each other than for epoch-pairings across topics (random permutation test: $p < .001$; Cohen's $d = 0.78$). Same result was obtained for all but three participants' individual FC matrices ($n = 23$, all $ps < .004$, $ds > 0.052$). Within-topic similarity was also significantly higher than across-topic similarity for each of the four topics ($ps < .001$; $ds > 0.53$; see Fig. 2C).

2) We also detected the subset of edges whose similarity was higher within- than across -topic: The relative contribution of each edge to the difference found for the group-average matrices (Fig. 3A) was evaluated by repeating the similarity analysis for each edge and applying FDR correction ($q = .05$) to control for the large number of statistical tests. $N = 615$ (32.52%) edges showed significant difference for this contrast in the alpha band.
In order to estimate the smallest network corresponding to topic sensitivity, the edges were ordered according to their effect size (d; Fig. 3B) and an Error-Correcting-Output-Codes classifier with support vector machine templates was trained on the FC data with the topic label as the output and the number of top contributing edges (n) as the step variable. Classifier performance at each step was evaluated by the classification error rate (CER) with five-fold cross-validation using 100 repetitions (Fig. 3C). After smoothing mean CER values as a function of edges, a cutoff was applied at the first local minimum, balancing complexity with performance. The final network was then composed of $n = 29$ edges and the final mean CER is 17.33% (Fig. 3D). Based on measures of node centrality (strength, degree, betweenness, and closeness), the most influential nodes of the network are in the left supramarginal, superior parietal, and cuneal areas, as well as in the right inferior parietal gyrus.

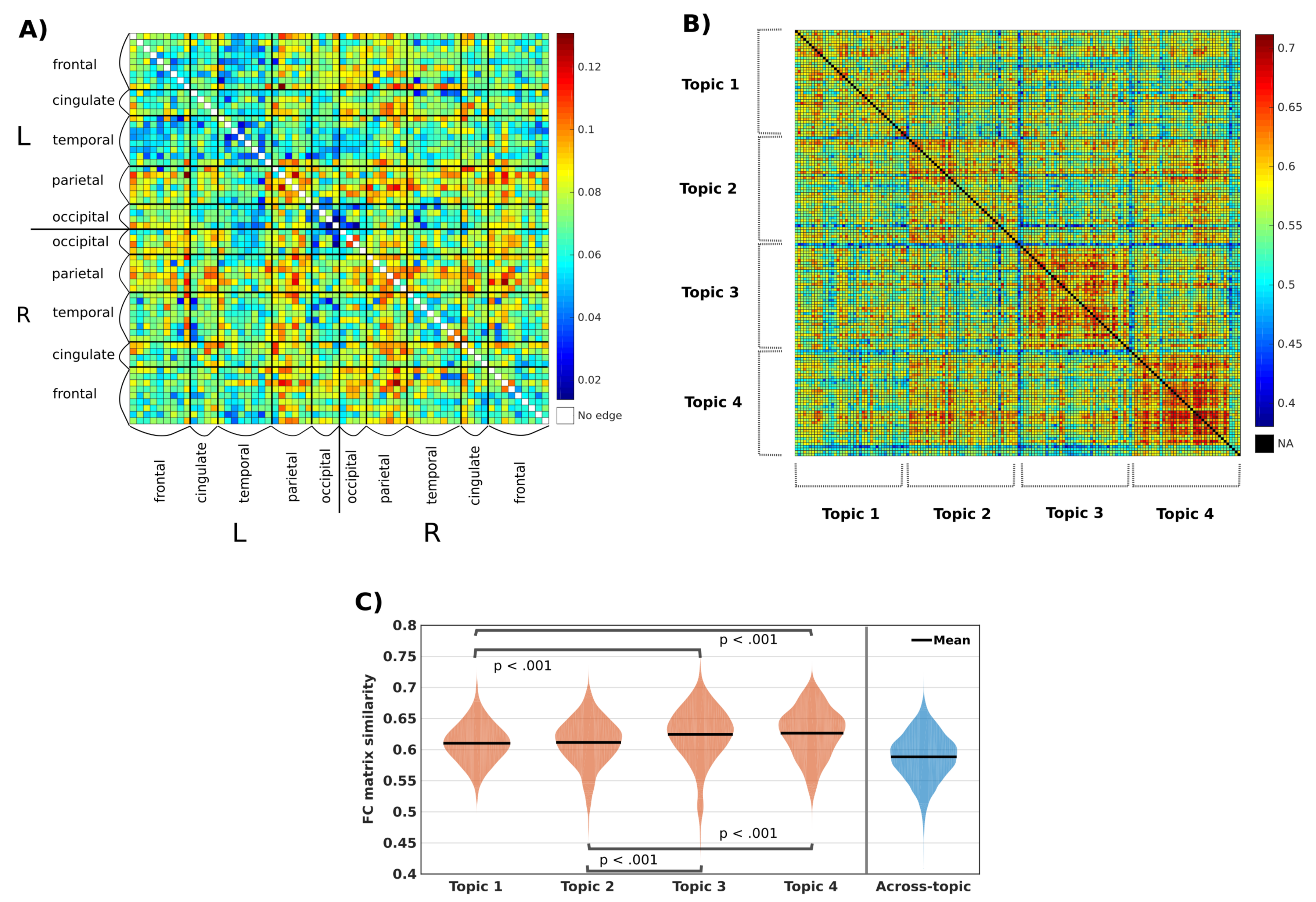

Fig. 2. FC matrix similarity within- and across-topics (alpha band). *(***A***)* Group average ($N = 26$) FC matrix averaged over all epochs (4 topics × 40 epochs = 160 epochs). Rows and columns represent the 62 ROIs grouped into larger anatomical regions, each cell representing the edge connecting the ROIs of the corresponding column and row. The matrix is symmetric to the diagonal. Color depicts connection strength. Cells with a connection strength not reaching the threshold in any of the participants are marked white. (**B**) Group average FC matrix similarity values for the four topics. Rows and columns represent epochs, grouped by topic and shown in their order within the speech recording. This matrix is also symmetric to the diagonal. Blue lines at the borders of the topics correspond to the first epoch of the recording. Blocks including the diagonal show the within-topic epoch-pairings. Color depicts similarity strength. (**C**) Estimated distributions and means of the within-topic FC matrix similarity values, separately for the four topics (pooled from all epoch-pairs of the given topic) and the distribution and means of across-topic FC matrix similarity values (pooled from all epoch-pairs belonging to two different topics). Each within-topic similarity sample differed from the across-topic similarity sample by at least $p < .001$. Bonferroni-corrected significant differences between the four within-topic similarities are marked on the figure.

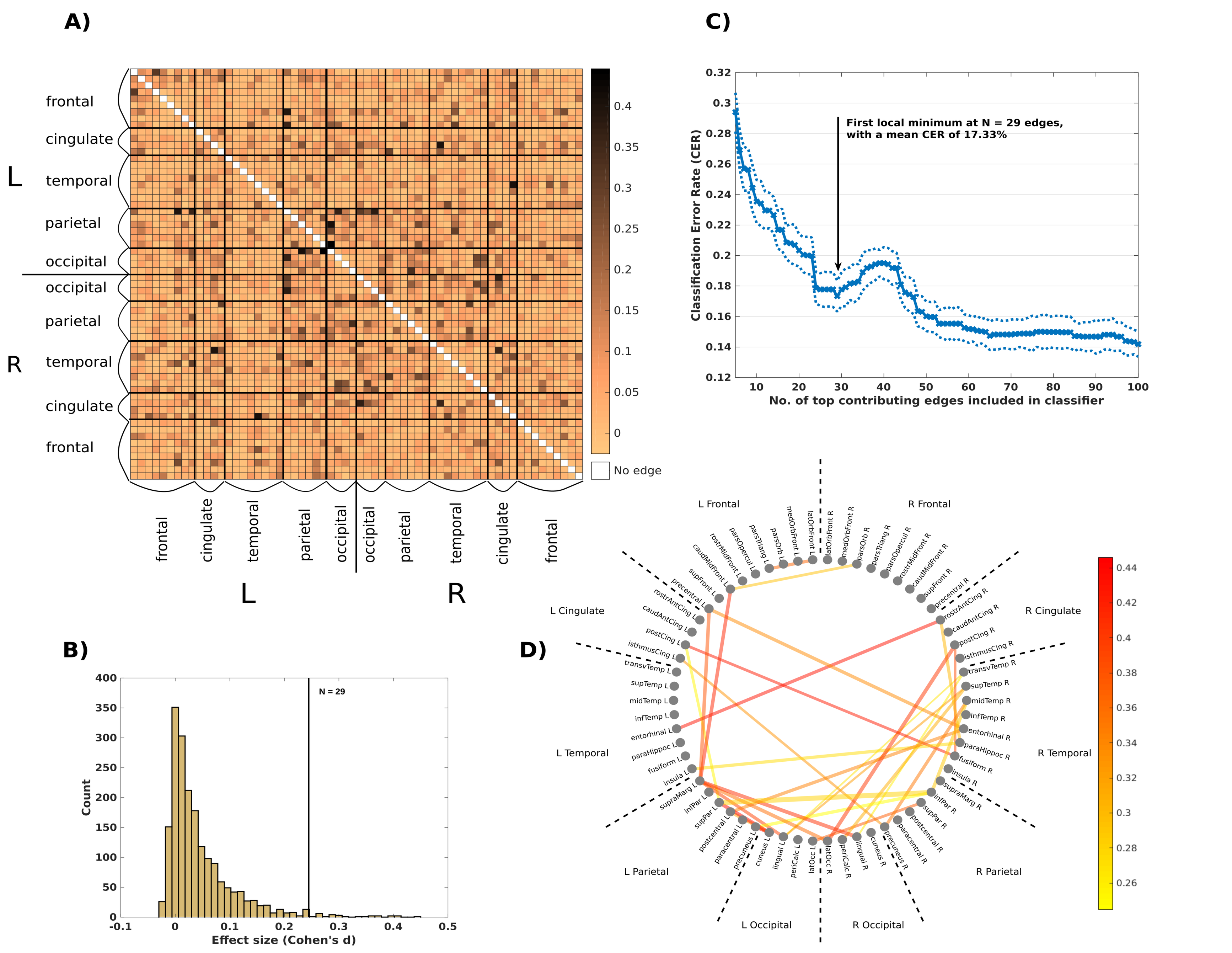

Fig. 3. Edge contributions to the FC similarity difference effect (alpha band). (**A**) Group-averaged (N=26) contribution (effect size, Cohen's d) of each edge to the group average FC similarity difference between within- and across-topic epoch-pairings, pooled across topics. Rows and columns represent the 62 ROIs grouped into larger anatomical regions. Color scale depicts edge contribution in terms of effect size (d). (**B**) Histogram of all effect sizes. Vertical line marks the cutoff point ($d = 0.247$) for the network of edges showing the largest sensitivity ($n = 29$). (**C**) Mean CER as a function of the number of edges included in the classifier. The arrow marks the cutoff at $n = 29$, that is, the first point after which CER increases with more edges added (first local minimum on median filtered CER values). Dotted lines depict the SD of CER values. (**D**) The 29 edges in the alpha band maximizing sensitivity to the similarity difference between within- and across-topics shown on a circle grouped by anatomical areas. Edge width corresponds to group-level connectivity strength and edge color marks relative contribution to the FC similarity difference between within- and across-topic epochs.

Conclusions

1) With linguistic and situational variables kept equal, **the brain networks extracted from EEG segments recorded while participants listened to spoken newspaper-like articles were significantly more like each other within than across different articles**. This relationship was found not only at the group level, but also for ca. 90% of the listeners, individually, as well as separately for each different topic.

2) It remains an open question if a representation of the topic is separate from the "lexicon", i.e., the representations assigning meaning to words. We would argue that our results are more consistent with the view that context is represented within the lexicon – e.g., word meanings are encoded in relation to their topic.

References and an online version of this poster are available at:
www.github.com/dharmatarha/SAMBA2022_poster

This research was supported by the National Research, Development and Innovation Office (grant K132642 awarded to I.W.; grant ANN131305 awarded to B.T.).

National Research, Development and Innovation Office