Functional geometry of auditory cortical resting state networks derived from intracranial electrophysiology

Matthew I. Banks[1,2*], Bryan M. Krause[1], D. Graham Berger[1], Declan I. Campbell[1], Aaron D. Boes[3], Joel E. Bruss[3], Christopher K. Kovach[4], Hiroto Kawasaki[4], Mitchell Steinschneider[5,6], Kirill V. Nourski[4,7]

[1]Department of Anesthesiology, University of Wisconsin, Madison, WI, USA
[2]Department of Neuroscience, University of Wisconsin, Madison, WI, USA
[3]Department of Neurology, The University of Iowa, Iowa City, IA 52242, USA
[4]Department of Neurosurgery, The University of Iowa, Iowa City, IA 52242, USA
[5]Department of Neurology, Albert Einstein College of Medicine, New York, NY 10461, USA
[6]Department of Neuroscience, Albert Einstein College of Medicine, New York, NY 10461, USA
[7]Iowa Neuroscience Institute, The University of Iowa, Iowa City, IA 52242, USA

***Corresponding author:**

Matthew I. Banks, Ph.D.
Professor
Department of Anesthesiology
University of Wisconsin
1300 University Avenue, Room 4605
Madison, WI 53706
Tel.: (608)261-1143
E-mail: mibanks@wisc.edu

Keywords:

Functional connectivity, hierarchy, fMRI, diffusion map embedding, electrocorticography, iEEG, ECoG

Acknowledgements

This work was supported by the National Institutes of Health (grant numbers R01-DC04290, R01-GM109086, S10OD025025, UL1-RR024979). We are grateful to Jess Banks, Alex Billig, Haiming Chen, Phillip Gander, Christopher Garcia, Matthew Howard, Ariane Rhone, and Matthew Sutterer for help with data collection, analysis, and comments on the manuscript.

Abstract

Critical details remain unresolved about the organization of the human auditory cortical hierarchy and its relationship to higher order brain networks. We investigated this organization using diffusion map embedding (DME) applied to resting state intracranial electroencephalography (iEEG) obtained in neurosurgical patients. DME was applied to functional connectivity measured between regions of interest (ROIs). ROIs exhibited a hierarchical organization, symmetric between the two hemispheres and robust to the choice of iEEG frequency band, connectivity metric, and imaging modality. Tight clusters of canonical auditory and prefrontal ROIs were maximally segregated in embedding space. Clusters consistent with ventral and dorsal auditory processing streams were paralleled by a cluster suggestive of a third stream linking auditory and limbic structures. Portions of anterior temporal cortex were characterized as global hubs. This approach lays the foundation for identifying network changes during active speech and language processing and elucidating mechanisms underlying disorders of auditory processing.

Introduction

Extensive research has contributed to our understanding of the organization of the human auditory cortical hierarchy[1,2], yet fundamental questions remain unresolved. These include identification of secondary, tertiary, and higher order regions, hindering elucidation of the neural basis of speech and language comprehension. For example, the anterior portion of the superior temporal gyrus (STGA) and planum polare (PP) are adjacent to auditory cortex on Heschl's gyrus, yet diverge from it functionally[3,4]. Conversely, the posterior insula (InsP) has response properties similar to those of core auditory cortex in the posteromedial portion of Heschl's gyrus (HGPM), yet is not considered a canonical auditory area[5]. Additionally, the upper and lower banks of the superior temporal sulcus (STSU, STSL) have distinct roles in speech and language processing[6,7], yet their functional relationships with other auditory areas are unclear.

Questions remain regarding broader organizational features as well. While the auditory hierarchy is posited to be organized along two processing streams[8-10], the specific brain regions involved and the functional relationships within each stream are vigorously debated[11-13]. Furthermore, communication between auditory cortex and hippocampus, amygdala, and anterior insula (InsA)[14] – areas involved in auditory working memory and processing of emotional aspects of auditory information[15-18] – suggests a third "limbic" auditory processing stream, complementary to the dorsal and ventral streams. Additionally, while hemispheric lateralization of speech and language processing is a widely accepted organizational feature[19,20], the degree to which it shapes the auditory hierarchy and is reflected in hemisphere-specific connectivity profiles is unknown[8,12,21-24].

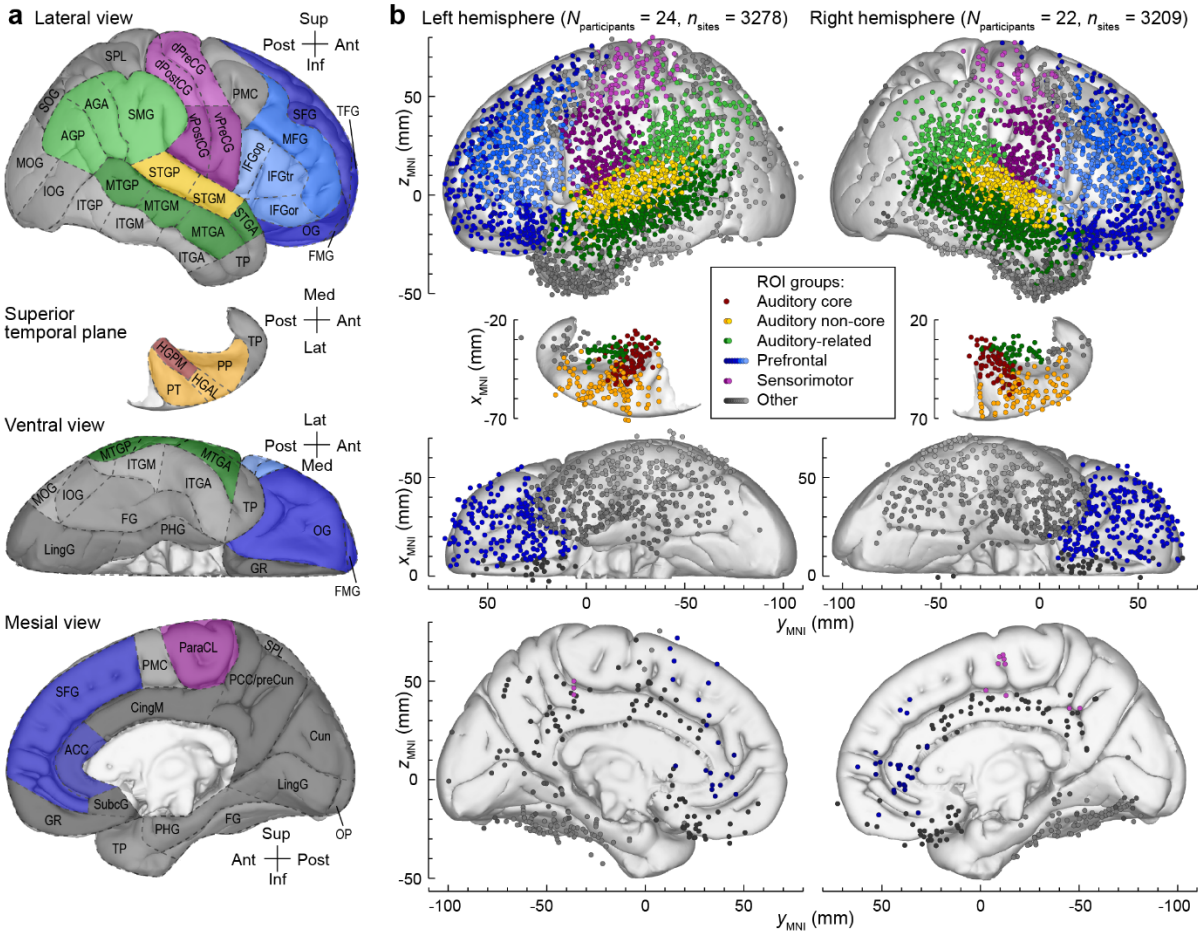
This study examined the organization of the human auditory cortical hierarchy within the broader context of cortical networks using resting state (RS) intracranial electroencephalography (iEEG) in neurosurgical patients. Although previous studies have investigated connectivity within the auditory cortical hierarchy using magnetic resonance imaging (MRI)[6,25,26], iEEG offers superior spatio-temporal resolution and is free of methodological problems that affect MRI in key regions such as the anterior temporal lobe[27,28]. Functional connectivity was measured as gamma band (30-70 Hz) power envelope correlations[29] between regions of interest (ROIs) based on an auditory-centric cortical parcellation scheme. Functional geometry of the network was defined based on the connectivity profiles of its constituent ROIs using diffusion map embedding (DME)[30,31]. DME maps ROIs into a Euclidean space where proximity of two ROIs reflects similarity between their connectivity profiles. DME has been previously applied to functional MRI (fMRI) data to demonstrate the hierarchical structure of canonical RS networks[32]. Here, we applied DME to iEEG data for the first time to gain new insights into the functional organization of human auditory cortical networks.

Results

DME applied to iEEG data

Intracranial electrodes densely sampled cortical structures involved in auditory processing in the temporal and parietal lobes, as well as prefrontal, sensorimotor, and other ROIs in 46 participants (a total of 6487 sites; Fig. 1, Supplementary Tables 1, 2). Electrode coverage was largely restricted to a single hemisphere in individual participants. On average, each participant contributed 141±54 recording sites, representing 28±7.9 ROIs (mean ± standard deviation) (see example in Fig. 2a). DME was applied to pairwise functional connectivity computed between recording sites in each participant. The functional connectivity matrix was normalized and thresholded to yield a diffusion matrix $\mathbf{P}_{\mathbf{symm}}$ with an apparent community structure along the horizontal and vertical dimensions (Fig. 2b). DME reveals the functional geometry of the sampled cortical sites by mapping the data of $\mathbf{P}_{\mathbf{symm}}$ into a low-dimensional embedding space where proximity between nodes represents similarity in their connectivity to the rest of the network (Fig. 2c; see Supplementary Fig. 1 for additional views).

DME exhibited superior signal-to noise characteristics compared to direct analysis of functional connectivity in 40 out of 46 participants (Supplementary Fig. 2). Sites that belonged to specific ROIs distributed to specific locations in embedding space. For example, in Figure 2c, tight clusters of auditory cortical sites (red/orange/yellow) and sites in prefrontal cortex (blue) were maximally segregated along dimension 1 (see Fig. 1 and Supplementary Table 3 for the list of abbreviations). Other ROIs [e.g., posterior portion of middle temporal gyrus (MTGP)] had a more distributed representation, consistent with their functional heterogeneity.



c

Auditory core	Heschl's gyrus, posteromedial portion	HGPM
Auditory non-core	Heschl's gyrus, anterolateral portion	HGAL
	Planum temporale	PT
	Planum polare	PP
	Superior temporal gyrus, posterior portion	STGP
	Superior temporal gyrus, middle portion	STGM
Auditory-related	Superior temporal gyrus, anterior portion	STGA
	Posterior insula	InsP*
	Superior temporal sulcus, upper bank	STSU*
	Superior temporal sulcus, lower bank	STSL*
	Middle temporal gyrus, posterior portion	MTGP
	Middle temporal gyrus, middle portion	MTGM
	Middle temporal gyrus, anterior portion	MTGA
	Supramarginal gyrus	SMG
	Angular gyrus, posterior portion	AGP
	Angular gyrus, anterior portion	AGA
Prefrontal	Inferior frontal gyrus, pars opercularis	IFGop
	Inferior frontal gyrus, pars triangularis	IFGtr
	Inferior frontal gyrus, pars orbitalis	IFGor
	Middle frontal gyrus	MFG
	Superior frontal gyrus	SFG
	Anterior cingulate cortex	ACC
	Orbital gyri	OG
	Transverse frontopolar gyrus	TFG
	Frontomarginal gyrus	FMG
Sensorimotor	Precentral gyrus, dorsal portion	dPreCG
	Postcentral gyrus, dorsal portion	dPostCG
	Precentral gyrus, ventral portion	vPreCG
	Postcentral gyrus, ventral portion	vPostCG
	Paracentral lobule	ParaCL

Other	Premotor cortex	PMC
	Cingulate gyrus, middle portion	CingM
	Cingulate gyrus, posterior portion / Precuneus	PCC/preCun
	Parahippocampal gyrus	PHG
	Fusiform gyrus	FG
	Inferior temporal gyrus, posterior portion	ITGP
	Inferior temporal gyrus, middle portion	ITGM
	Inferior temporal gyrus, anterior portion	ITGA
	Temporal pole	TP
	Anterior insula	InsA*
	Frontal operculum	fOperc*
	Parietal operculum	pOperc*
	Gyrus rectus	GR
	Subcallosal gyrus	SubcG
	Superior parietal lobule / Intraparietal sulcus	SPL
	Cuneus	Cun
	Lingual gyrus	LingG
	Occipital pole	OP
	Superior occipital gyrus	SOG
	Middle occipital gyrus	MOG
	Inferior occipital gyrus	IOG
	Amygdala	Amyg*
	Hippocampus	Hipp*
	Putamen	Put*
	Globus pallidus	GP*
	Caudate nucleus	Caud*
	Substantia innominata	SubInn*
	Ventral striatum	vStr*

*not pictured

Figure 1. ROIs and electrode coverage in all 46 participants. **a:** ROI parcellation scheme. **b:** Locations of recording sites, color-coded according to the ROI group, are plotted in Montreal Neurological Institute (MNI) coordinate space and projected onto the Freesurfer average template brain for spatial reference. Color shades represent different ROIs within a group. Projections are shown on the lateral, top-down (superior temporal plane), ventral and mesial views (top to bottom). Recording sites over orbital, transverse frontopolar, inferior temporal gyrus and temporal pole are shown in both the lateral and the ventral view. Sites in fusiform, lingual, parahippocampal gyrus and gyrus rectus are shown in both the ventral and medial view. Sites in the frontal operculum (n = 21), parietal operculum (n = 15), amygdala (n = 73), hippocampus (n = 77), putamen (n = 13), globus pallidus (n = 1), caudate nucleus (n = 10), substantia innominata (n = 5), and ventral striatum (n = 2) are not shown. See Supplementary Table 2 for detailed information on electrode coverage. **c:** ROI groups, ROIs and abbreviations used in the present study. See Supplementary Table 3 for alphabetized list of abbreviations.

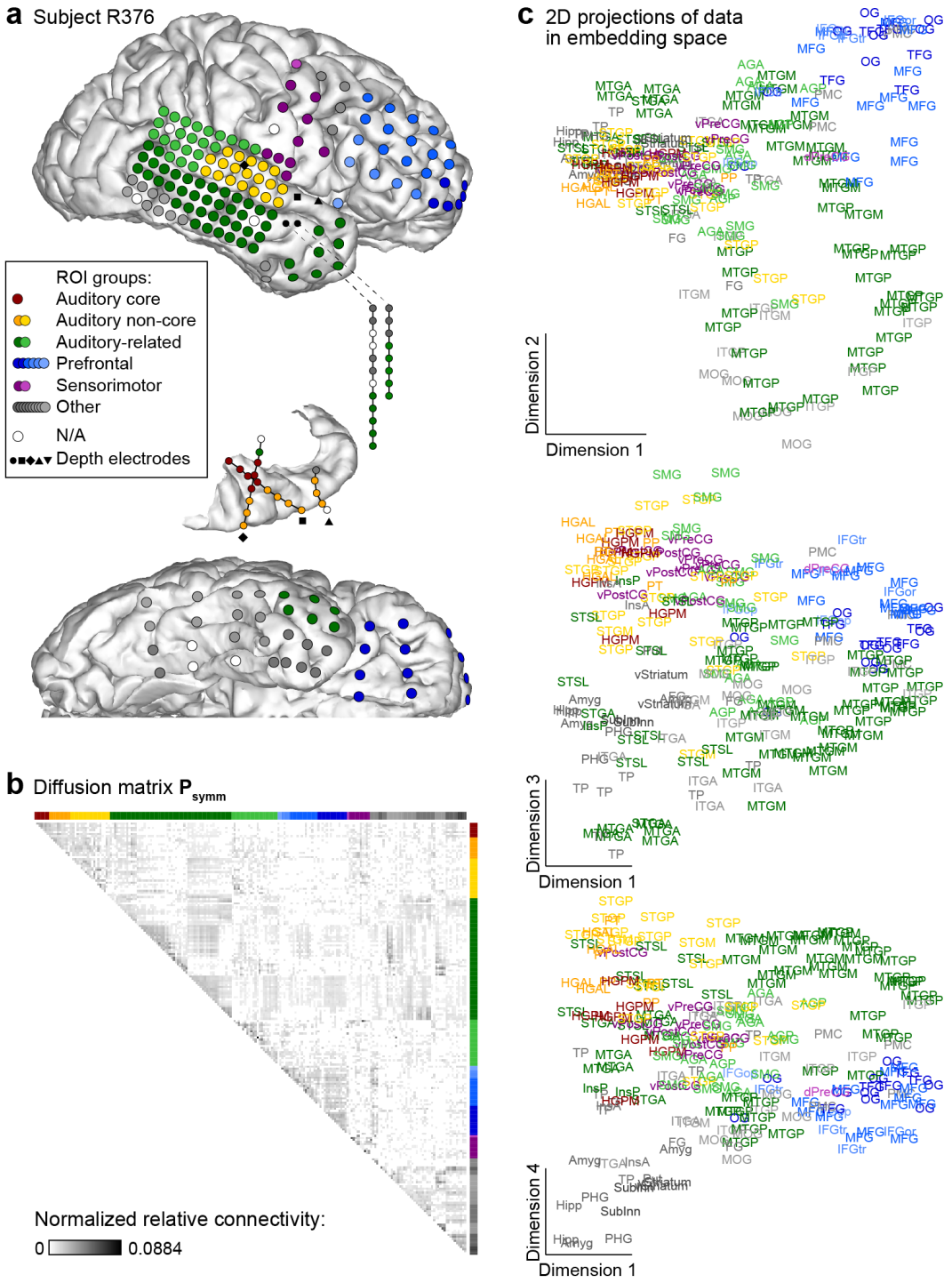


Figure 2. Functional geometry of cortical networks revealed by DME in a single participant (R376). **a:** Electrode coverage. **b:** Diffusion matrix $\mathbf{P_{symm}}$. **c:** Data plotted in the 1st and 2nd, 1st and 3rd, and 1st and 4th dimensions of embedding space (top to bottom). Two points that are close in embedding space are similarly connected to the rest of the network, and thus assumed to be functionally similar. Scale bar: 0.03.

126 *Functional geometry of cortical networks*

127 To pool data across participants with variable electrode coverage, $\mathbf{P}_{\mathbf{symm}}$ matrices were
128 computed at the ROI level and averaged across participants (Fig. 3a). The eigenvalue spectrum
129 $|\lambda_i|$ of $\mathbf{P}_{\mathbf{symm}}$ featured an inflection point for i>4 (Fig. 3a, inset), indicating that the first four
130 dimensions of embedding space accounted for the community structure of the data. The data
131 are plotted in the first four dimensions of embedding space in Figure 3b (see also
132 Supplementary Fig. 3 and Supplementary Movies 1 and 2), providing a graphical representation
133 of the functional geometry of all sampled brain regions. Functionally related ROIs clustered
134 together, and these clusters segregated within embedding space. For example, auditory cortical
135 and prefrontal ROIs were at opposite ends of dimension 1, and parietal and limbic ROIs were at
136 opposite ends of dimension 2. By contrast, some ROIs [e.g., STGA, anterior and middle portions
137 of middle temporal gyrus (MTGA, MTGM)] were situated in the interior of the data cloud.

138 The connectivity metric employed here discards components exactly in phase between two
139 brain regions, mitigating the influence of volume conduction[29]. However, brain areas that are
140 anatomically close to each other are often densely interconnected[33-37]. Thus, anatomical
141 proximity can contribute to the observed functional geometry. Auditory cortical ROIs followed
142 this rule and clustered together with one notable exception. PP, located immediately anterior
143 to anterolateral Heschl's gyrus (HGAL), segregated from the rest of auditory cortical ROIs along
144 dimension 2 in embedding space (Fig. 3b, upper panel, lower left corner). Separation in
145 embedding space was also observed along specific dimensions between other anatomically
146 adjacent ROIs, including STGA and STGM, temporal pole (TP) and the rest of the anterior
147 temporal lobe (ATL), and InsA and InsP. Anatomical proximity explained 15% of the variance in
148 embedding distance (mean adjusted r^2 = 0.15 for regressions between anatomical and
149 embedding Euclidean distance, calculated separately for each ROI). Thus, the embedding
150 representation elucidates organizational features of ROIs beyond anatomical proximity.

151 The grouping of canonical auditory ROIs is apparent in Figure 3b, as PT, HGAL, and middle and
152 posterior portions of the superior temporal gyrus (STGM, STGP) were all close to HGPM in
153 embedding space. Interestingly, so was STSU, which was significantly closer to auditory cortex
154 in embedding space compared to STSL (p=0.002). This separation between STSL and STSU is
155 consistent with differences in their response properties reported recently[7]. Particularly,
156 responses in STSL, but not STSU, were predictive of performance in a semantic categorization
157 task. This suggests that in embedding space, STSL would be closer to regions involved in
158 semantic processing compared to STSU. Indeed, a permutation analysis revealed that STSL was
159 closer to ROIs reported to contribute to semantic processing [inferior frontal gyrus (IFG) pars
160 operculum/triangularis/orbitalis (IFGop, IFGtri, IFGor), TP, STGA, MTGA, MTGP, anterior and
161 posterior portions of inferior temporal gyrus (ITGA, ITGP), anterior and posterior angular gyrus
162 (AGA, AGP), supramarginal gyrus (SMG)][38-40] compared to STSU (p=0.010).

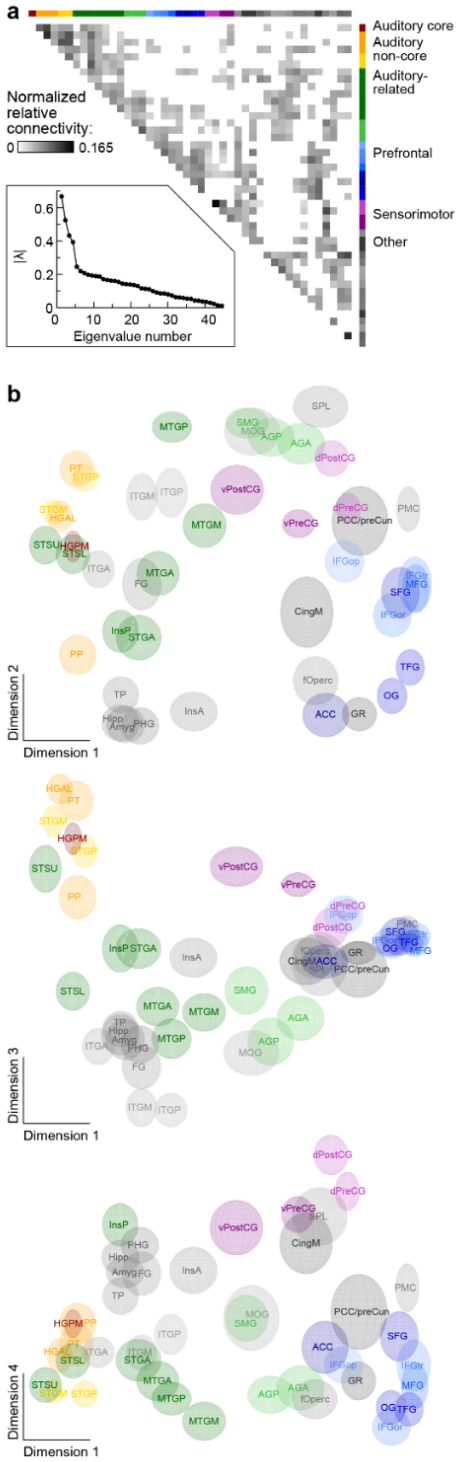


Figure 3. Summary of functional geometry of cortical networks via DME. **a**: Average diffusion matrix. **Inset:** Eigenvalue spectrum. **b:** Data plotted in the 1st and 2nd, 1st and 3rd, and 1st and 4th dimensions of embedding space (top to bottom). Variance estimates on the locations of each ROI in embedding space were obtained via bootstrapping and are represented by the size of the ellipsoid for each ROI. Scale bar: 0.1.

169 InsP responds robustly to acoustic stimuli[5], suggesting that a portion of this area could be
170 considered an auditory region[41]. It can track relatively fast (>100 Hz) temporal modulations,
171 similar to HGPM[5,42], possibly due to direct inputs from the auditory thalamus. However, InsP
172 was functionally segregated from HGPM and was situated between auditory and limbic ROIs,
173 consistent with the broader role of InsP in polysensory exteroceptive processing and
174 interoception[43,44].

175 Figure 3b also characterizes the temporal and parietal ROIs outside auditory cortex that are
176 nonetheless part of the extended auditory network, including components of the dorsal and
177 ventral processing streams. These 'auditory-related' ROIs (shades of green in Fig. 3b), were
178 distributed along a considerable extent of all four dimensions, consistent with functional
179 heterogeneity of these regions and their involvement in multimodal integration[45].

180 So far, the results of analyses have been presented at a single spatial scale (t = 1), which
181 emphasized local network structure. Analysis of data at a coarser spatial resolution (t = 5)
182 showed that the dominant structure in the data was governed by dimension 1, with auditory
183 ROIs at one end and prefrontal ROIs at the other (Supplementary Fig. 4). This indicates that the
184 functional distance between these two groups of ROIs is the most prominent of all
185 organizational features in the data.

186

187 *Hierarchical clustering of ROIs*

188 Hierarchical clustering applied to the first four dimensions of the embedded data revealed
189 segregation of the ROIs into several well-delineated clusters (Fig. 4). Auditory cortical ROIs
190 (excluding PP) formed an 'Auditory' cluster with STSU. Another major cluster (labeled 'Limbic')
191 included ROIs traditionally considered part of the limbic system [parahippocampal gyrus (PHG),
192 amygdala and hippocampus], ATL ROIs (TP, STGA, PP), and the insula. ROIs typically considered
193 part of the ventral and dorsal auditory streams segregated into two clusters. Additional clusters
194 included ROIs involved in executive and sensorimotor functions (labeled 'Executive' and
195 'Action', respectively in Fig. 4). Thus, the hierarchical clustering analysis revealed a segregation
196 of ROIs in embedding space that aligned with known functional differentiation of brain regions.
197 This analysis expands our understanding of functional relationships within the auditory cortical
198 hierarchy (e.g., clustering of STSU with canonical auditory ROIs and segregation of PP and InsP
199 from HGPM) and between auditory and higher order cortical areas (e.g., the proximity of
200 'Auditory' and 'Limbic' clusters).

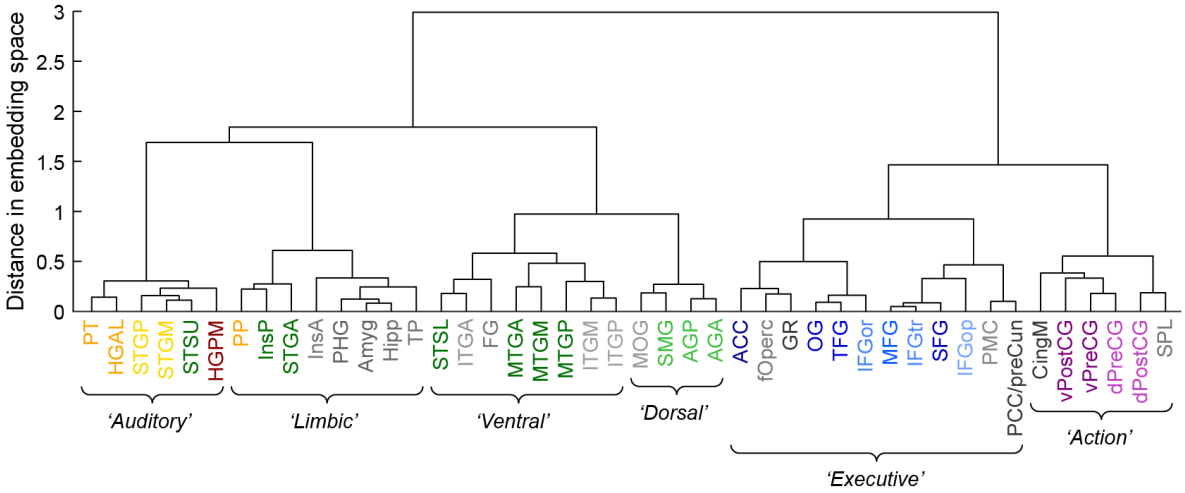


Figure 4. Hierarchical clustering of embedding data. Dendrogram showing linkages between ROI groups identified using agglomerative clustering.

Identification of network hubs

Identification of 'global hubs' within brain networks is critical for understanding their topology[46]. These nodes integrate and regulate information flow in the network by virtue of their centrality and strong connectivity. DME can identify global hubs, as the closer an ROI is to the center of the data cloud in embedding space, the more equal is its connectivity to the rest of the network. This is illustrated in Figure 5a, which depicts a simulated network of five ROIs, with one serving as a global hub (Fig. 5a, left panel, green). The network structure can also be represented as an adjacency matrix, wherein the hub ROI has strong connectivity with other ROIs (Fig. 5a, middle panel). In embedding space, this ROI occupies a central location, with the other four serving as spokes, i.e., nodes that interact with each other through the central hub (Fig. 5a, right panel).

However, a node's proximity to the center of the data cloud reflects the homogeneity of its connectivity to the rest of the network. Thus, a node can appear at a central location if it is weakly but consistently connected to all other nodes. By contrast, a node strongly connected to a small number of nodes ('local hub') can have high mean connectivity but still be distant from the center of the data cloud. Combining distance from the center with mean connectivity (Fig. 5b) allows for classification of ROIs as global versus local hubs and for distinguishing hubs from spokes and central nodes that have relatively weak overall connectivity ('bridge'). ATL ROIs (MTGA, STGA, but not TP) and MTGM lie in the upper left quadrant of the plot >2 standard deviations from the center of the data cloud (outer dashed ellipse) and appear as global hubs. ITGA, posterior cingulate/precuneus (PCC/preCun), and STSL also exhibited hub-like properties, i.e., were located in the upper left quadrant of Figure 5b.

STSU and orbitofrontal gyrus (OG) tended toward high connectivity but were located at a greater distance from the center and can thus be considered local hubs. Primary sensory and motor regions and canonical auditory cortex (except PP) tended to behave as spokes, occupying the lower right quadrant of the graph. Finally, the ROIs in the lower left quadrant can be considered bridges between isolated clusters of ROIs, characterized by more selective connectivity patterns. Thus, DME can identify topological features critical to information flow within cortical networks.

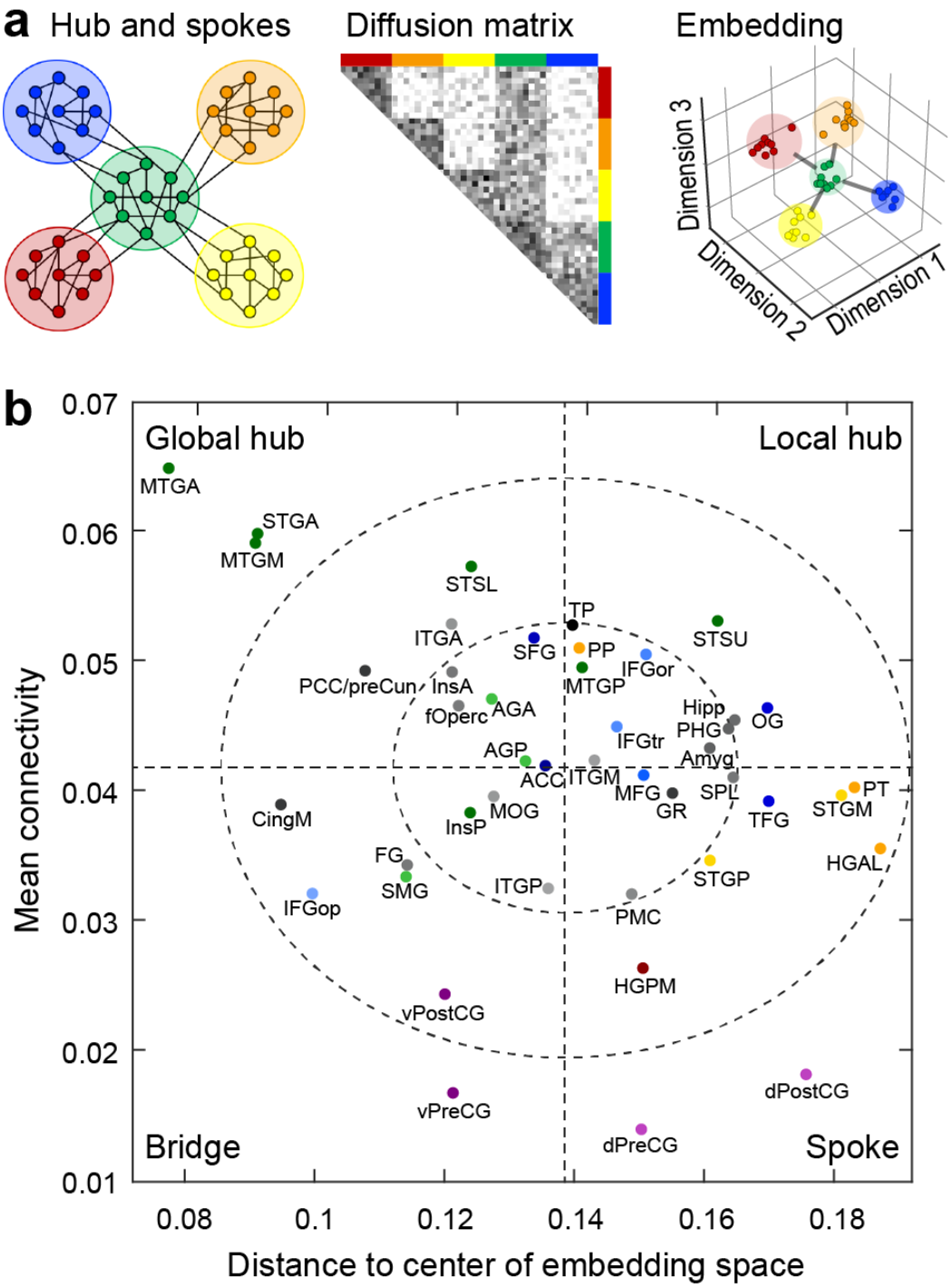


Figure 5. Identification of network hubs. **a:** Schematic example illustrating the central positioning of global hubs in embedding space. **b:** ROIs from average embedding are plotted according to their mean connectivity to the rest of the network versus their distance to the centroid of the data cloud in the first four dimensions of embedding space. Dashed lines denote across-ROI means. Dashed ellipses represent 1 and 2 standard deviations from the mean.

Comparisons across hemispheres

Speech and language networks are lateralized in the human brain, with nearly all right-handed and most left-handed individuals left hemisphere language-dominant[47]. However, both hemispheres are activated during speech processing[9,22,48,49], and the extent to which lateralization is reflected in asymmetries in the organization of auditory networks is unclear. We addressed this issue by comparing the geometry of cortical networks derived from participants with electrode coverage in the language-dominant (N = 22) versus non-dominant (N = 21) hemisphere. ROIs exhibited a symmetric functional organization in embedding space (Supplementary Fig. 5). Pairwise inter-ROI distances in embedding space, calculated separately for dominant versus non-dominant hemisphere, were highly correlated (r = 0.77), with no obvious outliers (Fig. 6, left panel). Furthermore, this was true specifically for ROIs involved in speech and language comprehension and production [PT, PP, STSL, STGP, STGM, STGA, SMG, AGA, premotor cortex (PMC), precentral gyrus (PreCG), IFGop, IFGtr][12,50,51] (r = 0.71; Fig. 6, right panel). Permutation analysis indicated that neither the positions of ROIs in embedding space nor the pairwise distances between ROIs were significantly different between dominant and non-dominant hemispheres (all p-values > 0.05). Thus, hemispheric asymmetry of functional organization of speech and language networks was not detectable in RS connectivity.

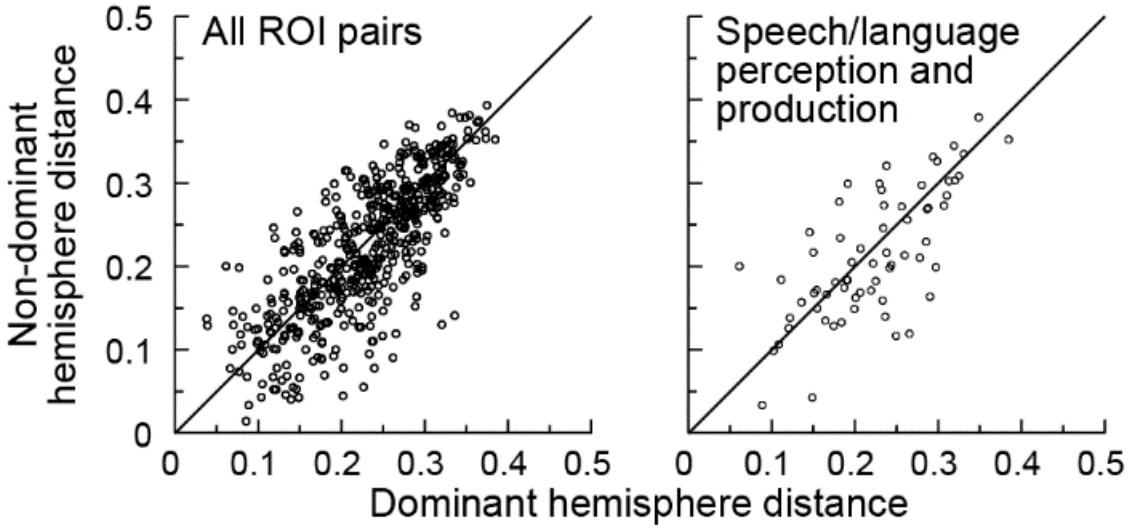


Figure 6. RS connectivity is symmetric between hemispheres. Inter-ROI distances in embedding space for non-dominant versus dominant hemisphere participants. Pairwise distances between all ROIs and between ROIs involved in speech and language perception and production (PT, PP, STSL, STGP, STGM, STGA, SMG, AGA, PMC, PreCG, IFGop, IFGtr) are shown in the left and right panel, respectively. Note that after splitting the data into the two subsets (dominant and non-dominant) STSU did not meet the inclusion criteria for analysis presented in the right panel (see Methods, Supplementary Table 2).

Stability of functional geometry across frequency band and connectivity measures

Embeddings presented so far have been derived from gamma-band power envelope correlations. Other bands (theta, alpha, beta, high gamma) and a different measure (debiased weighted phase lag index, wPLI[52]) produced similar embeddings. For power envelope correlations, inter-ROI distances were highly similar for adjacent bands ($r \geq 0.83$), and even for non-adjacent bands ($r \geq 0.72$; Supplementary Fig. 6a). For wPLI, correlations were only slightly weaker (0.59-0.79). Correlations were strong even across the two connectivity measures (0.58-0.78). Thus, DME identified overall, rather than band- or measure-specific, organizational features of cortical networks.

However, a particular band or measure might be preferred if it produced narrower estimation margins in the functional geometry. An overall relative uncertainty was calculated as the variation of ROI position across bootstrapped samples relative to the average distance to other ROIs. Relative uncertainty was similar across bands and measures but was lowest for power envelope correlations in beta, gamma, and high gamma bands (Supplementary Fig. 6b). These analyses suggest that gamma power envelope correlations are a robust measure for exploring functional geometry.

Comparison to embeddings derived from RS-fMRI data

Intracranial recordings sample the brain non-uniformly and sparsely as dictated by clinical considerations. To examine the impact of this sampling, DME was applied to RS-fMRI data available in a subset of ten participants. We first verified the consistency of functional geometry derived from the two modalities in the same participants (Fig. 7). Connectivity matrices were constructed based on RS-fMRI data from voxels located at iEEG recording sites and grouped into the same ROIs as in Figure 1. The iEEG and fMRI embeddings averaged across participants were qualitatively similar (Fig. 7a, b), and the overall organization derived from this subset was consistent with that observed in the full iEEG dataset (cf. Fig. 3b). Inter-ROI distances in the fMRI and iEEG embedding spaces were strongly correlated (Fig. 7), with highest correlations for beta-, gamma- and high gamma-band envelopes ($r > 0.5$; Fig. 7d, line and symbols). Similar results were observed at the individual participant level as well, albeit with smaller r values (Fig. 7d, box and whiskers plot).

The analysis presented in Figure 7 indicates that fMRI data can be used to address two questions regarding the effects of limited, non-uniform sampling. The first question is the effect of non-uniformly sampling only a subset of brain regions. We used a standard parcellation scheme developed for fMRI data (Schaefer-Yeo 400 ROIs;[53]) rather than the iEEG parcellation scheme introduced in Figure 1. For each participant, embeddings were derived from RS-fMRI connectivity matrices computed from all cortical ROIs ("Full fMRI", Fig. 8a, leftmost column). A subset of these ROIs ["Full fMRI (iEEG subset)", Fig. 8a, 2nd column] were retained for comparison to embeddings computed from the fMRI data corresponding to the ROIs sampled

by iEEG [“Partial fMRI (ROI level)”, Fig. 8a, 3rd column]. The two embeddings were compared by computing the correlation between inter-ROI distances in the respective embedding spaces (Fig. 8b). Although the scale of the embeddings was different for the full fMRI versus partial fMRI data (because the number of dimensions was different), the two were highly correlated ($r = 0.94$; Fig. 8c). Thus, embeddings constructed from the portion of the brain sampled by iEEG were nearly identical to embeddings derived from the whole brain.

The second question is the effect of representing an entire ROI by sparse sampling with a limited number of electrodes. To address this question, embeddings were computed from the averages across entire ROIs in each participant [“Partial fMRI (ROI level)”, Fig. 8a, 3rd column] and from averages of the fMRI voxels located at iEEG recording sites [“Partial fMRI (site level)”, Fig. 8a, rightmost column]. ROI- and site-level embedding distances were strongly correlated ($r > 0.5$; Fig. 8c). Thus, sparse sampling within an ROI had a greater impact on estimates of functional geometry than limited sampling of the complete set of ROIs. Overall, ROIs were faithfully represented in embedding space even when DME was based on a small number of locations within ROIs. Taken together, these results indicate broad consistency between functional organization derived from iEEG and fMRI and the robustness of this approach to sparse sampling afforded by iEEG recordings.

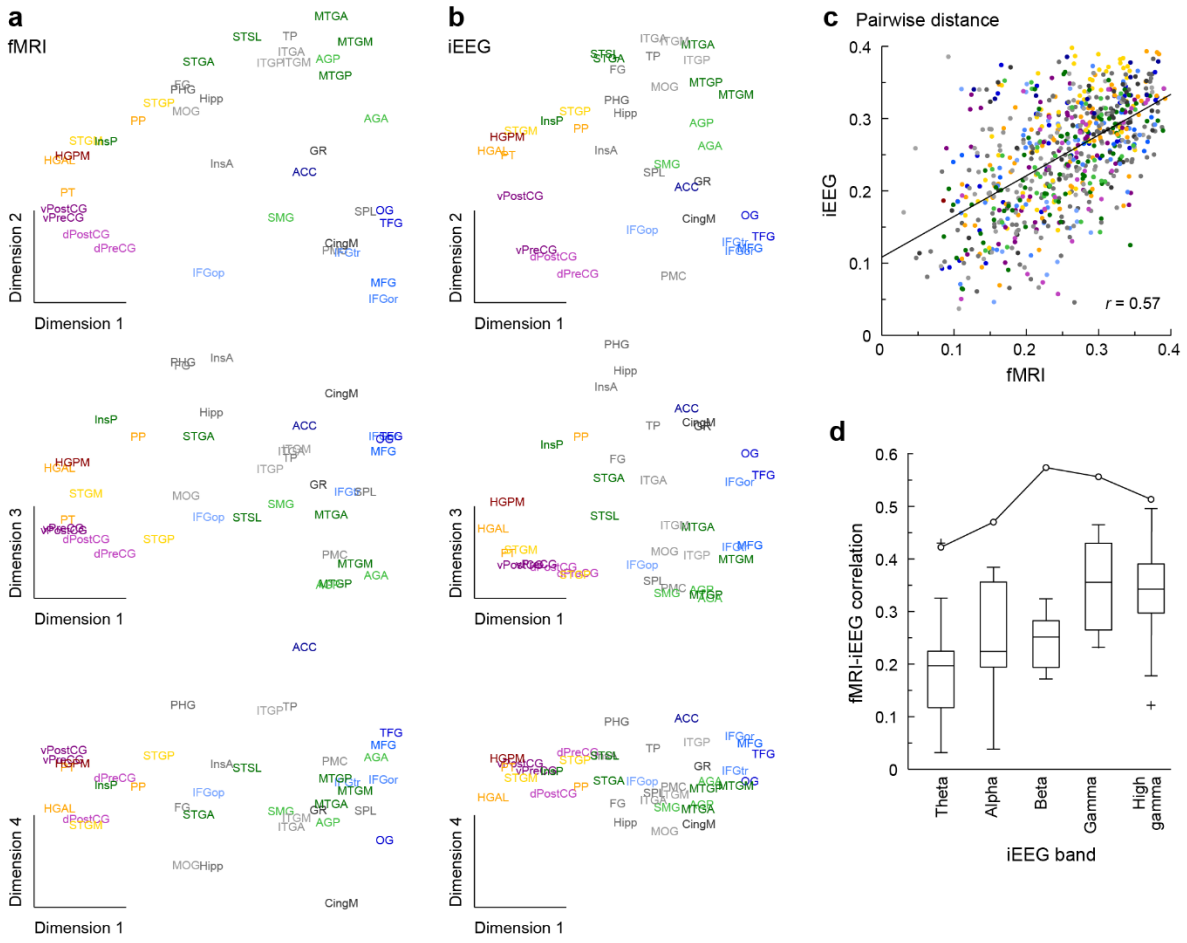


Figure 7. Comparison of iEEG and fMRI connectivity data in embedding space. **a:** Participant-averaged embeddings for iEEG (gamma band power envelope correlations). **b**: Participant-averaged embeddings for fMRI. Scale bar: 0.1. **c:** Inter-ROI embedding distances computed from the data in **a** and **b**. **d:** Summary of distance correlations at each frequency band for participant-averaged data (line and symbols) and single participants (box and whiskers). *t* = 1 for all embeddings.

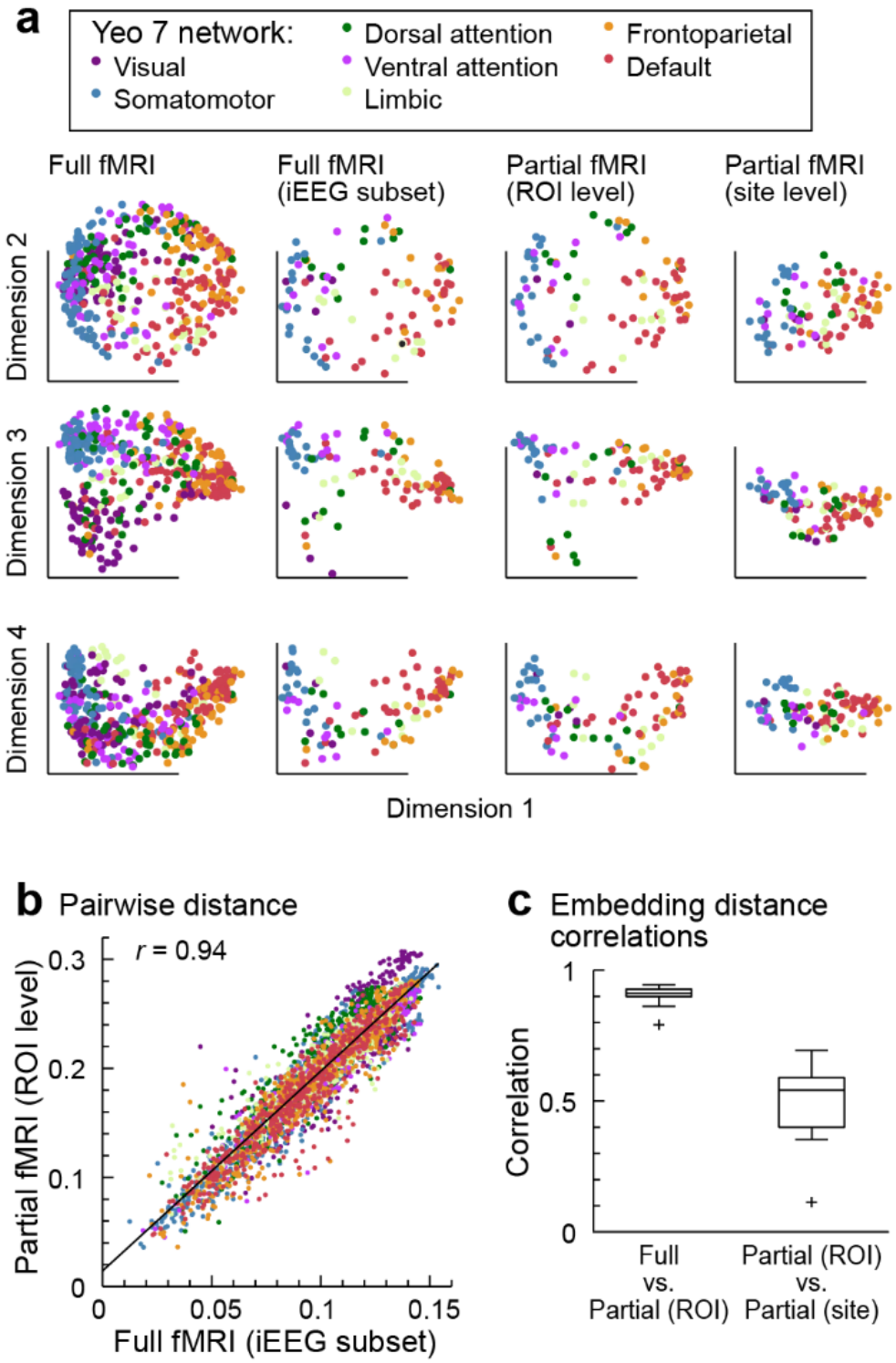


Figure 8. Effect of constructing embeddings from only a subset of ROIs in the brain. Scale bar: 0.1 (full fMRI), 0.2 (partial fMRI).

Discussion

Organization of auditory cortical networks

We have shown that DME applied to iEEG data can be used to characterize the organization of the human auditory cortical hierarchy. Previous work in macaque has defined over a dozen auditory cortical fields based on cytoarchitectonics, connectivity, and response properties[54]. By contrast, there is no consensus on how auditory cortex is organized in humans, with multiple candidate parcellations based on cytoarchitectonics, tonotopy or myeloarchitecture [55-58]. Our results contribute to this body of knowledge by showing that superior temporal ROIs including core auditory cortex (HGPM) and putative auditory belt and parabelt areas (PT, HGAL, STGP, STGM[55,58]) cluster together in embedding space, indicating functional similarity. We could also discern hierarchical relationships within this cluster by combining embedding analysis with mean connectivity. Specifically, despite the proximity of HGPM to other auditory ROIs in embedding space (Fig. 3b), it exhibited weaker mean connectivity, suggesting that it functions as a spoke rather than a local hub (Fig. 5b). This is consistent with the position of HGPM at the lowest level of the auditory cortical hierarchy. STGP had an intermediate position between HGPM and HGAL/PT/STGM in terms of its eccentricity and mean connectivity. This suggests that STGP is the locus of relatively early non-core auditory cortex[59,60], occupying a low-level position in the hierarchy[61].

Functional differentiation between STSU and STSL

The superior temporal sulcus is a critical node in speech and language networks linking canonical auditory cortex with higher order temporal, parietal, and frontal areas[6,48,50,62-64]. Previous studies have shown that STSU and STSL differ in cytoarchitecture[65] and have distinct responses to speech[23,66-68]. A recent iEEG study demonstrated enhanced, shorter-latency, responses to speech syllables in STSU compared to STSL[7]. STSU is traditionally not considered part of canonical auditory cortex (but see [56]), yet it clustered with auditory cortical ROIs. STSL, by contrast, was closer in embedding space to semantic ROIs. This is consistent with iEEG evidence that responses in STSL, but not STSU, correlated with performance on a semantic categorization task[7]. The regions specifically involved in semantic processing is a current topic of debate, with multiple competing models[28,38-40], and we defined a list of semantic ROIs by combining across these models. Using a family of plausible candidate ROIs for this comparison dilutes the importance of including or excluding any particular ROI. Taken together, the results firmly place STSU and STSL at different levels of the auditory cortical hierarchy.

361 *Functional and theoretical framework of a limbic auditory pathway*

362 Multiple lines of evidence support a pathway linking auditory cortical and limbic structures[69-72]
363 that subserves auditory memory[14,17,18] and affective sound processing[73]. The data presented
364 here contribute to our understanding of this pathway. Clustering analysis identified a set of
365 ROIs (PP, InsP, STGA, InsA, TP) positioned between auditory and limbic cortex (Fig. 4). PP is
366 anatomically close and connected to auditory cortex[74], yet it is unique in auditory-responsive
367 cortex for its syntactic-level language processing[3] and its preferential activation by music, which
368 has a strong affective component[4]. The recently reported connectivity between Hippocampus
369 and PP[75] is consistent with PP's role as an intermediate node in this stream. Similarly, InsP,
370 despite its anatomical proximity and overlapping response properties with HGPM, is likely
371 involved in the transformation of auditory information in auditory cortex to affective
372 representations in InsA[5].

373 The ATL structures STGA and TP are involved in semantic processing[3,28] and auditory memory[76],
374 in particular the representation and retrieval of memories for people, social language, and
375 behaviors ('social knowledge')[77]. Tight clustering of TP with limbic ROIs in embedding space is
376 consistent with its previously reported functional association with limbic cortex[78,79], with which
377 TP shares key features of laminar cytoarchitecture and strong connectivity[75]. We suggest that
378 the organization depicted in Figures 3 and 4, combined with evidence for bidirectional
379 information sharing between auditory cortex and limbic areas, merits the identification of a
380 third auditory processing stream alongside the dorsal and ventral streams[8,80]. This 'limbic
381 stream' would underlie auditory contributions to affective and episodic memory processing.

382

383 *Ventral and dorsal streams linking auditory and frontal cortex*

384 Current models of speech and language processing posit the existence of ventral and dorsal
385 processing streams linking non-core auditory cortex with PMC and inferior frontal gyrus via
386 several distinct anatomical pathways encompassing temporal, parietal, and frontal cortex [8-10,50].
387 Despite substantial experimental evidence supporting these models, there is a lack of
388 consensus on the specific functions subserved by the two streams. For example, the dorsal
389 stream has been envisioned to subserve spatial processing ("where"[8]), sensorimotor
390 integration ("how"[9]), and syntactic processing[10]. There is a parallel debate about the specific
391 cortical regions comprising the two streams.

392 As broadly predicted by these models, temporal and parietal ROIs segregated in embedding
393 space in the analysis presented here (Fig. 3b, 4). We observed a cluster that included STSL,
394 middle and inferior temporal gyrus ROIs, in conformity with the ventral auditory stream
395 proposed by Hickok and Poeppel[9] and Friederici[10]. By contrast, the cluster that included SMG,
396 AGP, and AGA aligned with the dorsal processing stream as proposed by Rauschecker and
397 Scott[8]. Association of FG and MOG with the ventral and dorsal clusters, respectively, likely
398 represents the sharing of information across sensory modalities.

A previous fMRI-based DME study found that primary sensory and default mode ROIs segregated along the first dimension in embedding space[32]. Coverage of mesial cortex in our dataset was limited, precluding a direct comparison. However, the striking separation between auditory and prefrontal cortex in embedding space shown here, and its robustness to the choice of the parameter *t*, indicate that the current results align well with the previous report. This separation places auditory and frontal regions at opposite ends of the auditory processing hierarchy, linked by ventral and dorsal processing streams[8-10].

Network hubs

Hubs in brain networks play a critical role in integrating distributed neural activity[46]. In the present analysis, global hubs were characterized by their central location within embedding space and high mean connectivity (Fig. 5). These hubs included STGA and MTGA, both components of the ATL. Previous reports indicate that ATL serves as a transmodal hub, transforming sensory domain-specific to domain-general representations[28,81,82] and playing a central role in semantic processing and social memory[28,77,83]. MTGM also appears as a global hub, even though it is not formally part of the ATL. Interestingly, patients with semantic dementia have ATL degeneration[84,85], but the damage is often more widespread and can include MTGM[86].

Several other ROIs also exhibited hub-like properties. These included ITGA (part of ATL), the well-established global hub PCC/preCun[32], and STSL. The latter observation is consistent with the diverse and multimodal functionality attributed to the STS previously[63,87].

Unlike other ATL structures, TP does not appear as a global hub in Figure 5b. The close association of TP with limbic structures in embedding space suggests that TP mediates interactions between multimodal integration centers in the ATL and structures subserving memory functions. More broadly, the heterogeneity of ATL ROIs in terms of their global hub-like connectivity profiles conforms to the observation that the terminal fields of white matter tracts converging in the ATL only partially overlap[28,88,89].

Hemispheric lateralization

Although speech and language networks are classically described as highly lateralized, imaging studies have demonstrated widespread bilateral activation during speech and language tasks[90-92]. We found no evidence for hemispheric differences in cortical functional organization based on analysis of all sampled brain regions, nor when comparison was restricted to speech and language ROIs (Fig. 6). These results are consistent with a recent fMRI study[26] showing similar RS connectivity patterns between left and right temporal lobe (but see[6]).

434 A recent study that applied DME to RS-fMRI data presented evidence for hemispheric
435 asymmetry in semantic networks, which correlated with semantic task performance[93]. The
436 large dataset in that study and the availability of data from both hemispheres in every
437 participant likely facilitated detection of those differences. Another possible reason for this
438 divergence is that we investigated hemispheric differences based on all relevant dimensions of
439 embedding space, whereas the previous study relied only on position along the first dimension.
440 In addition, hemispheric asymmetry in network connectivity increases along the cortical
441 hierarchy[94]. Thus, ROIs involved in heteromodal semantic processing would likely show greater
442 hemispheric differences compared to speech and language ROIs, which include lower order
443 sensory regions. The present study suggests that neither anatomical nor task-specific functional
444 hemispheric asymmetries necessarily translate into asymmetric network configurations during
445 resting state. This does not exclude the possibility of asymmetries emerging during auditory
446 tasks, for example reflecting hemispheric biases in spectral and temporal processing[9,12].

447

448 *Caveats & limitations*

449 A key concern regarding all human iEEG studies is that participants may not be representative
450 of a healthy population. In the present study, results were consistent across participants
451 despite differences in seizure disorder histories, medications, and seizure foci, and aligned with
452 results obtained previously in healthy participants[32]. Another caveat is that our dataset,
453 however extensive, did not sample the entire brain, and it was not possible to infer connectivity
454 with unsampled regions. To address this, we applied DME analysis to fMRI data to establish that
455 the organization of ROIs in embedding space was robust to the exclusion of unsampled ROIs.
456 Although there was a greater effect of sparse, non-uniform sampling within an ROI, there was
457 still considerable similarity in functional organization to embeddings derived from averages
458 across the entire ROI.

459 While subcortical structures (e.g., thalamus) that link sensory and higher order networks[95] were
460 not sampled, the functional organization presented here was likely influenced indirectly by
461 thalamo-cortical pathways[61,96]. Previous fMRI studies of RS networks focused exclusively on
462 cortical ROIs and did not consider the role of the thalamus and other subcortical structures.
463 Despite this limitation, these studies have yielded valuable insights into the functional
464 organization of the human cortical networks[97,98].

465

466 *Concluding remarks and future directions*

467 This study extends the DME approach to characterize functional relationships between cortical
468 regions investigated using iEEG recordings. These data help resolve several outstanding issues
469 regarding the functional organization of human auditory cortical networks and stress the
470 importance of a limbic pathway complementary to the dorsal and ventral streams. These

results lay the foundation for future work investigating network organization during active speech and language processing. While the current work focused on auditory cortical networks, this approach can be readily generalized to advance our understanding of changes in brain organization during sleep and anesthesia, disorders of consciousness, as well as reorganization of cortical functional geometry secondary to lesions.

Online Methods

Participants

The study was carried out in 46 neurosurgical patients (20 females) diagnosed with medically refractory epilepsy. The patients were undergoing chronic invasive electrophysiological monitoring to identify seizure foci prior to resection surgery (Supplementary Table 1). Research protocols were approved by the University of Iowa Institutional Review Board and the National Institutes of Health, and written informed consent was obtained from all participants. Research participation did not interfere with acquisition of clinically necessary data, and participants could rescind consent for research without interrupting their clinical management.

All participants except two were native English speakers. The participants were predominantly right-handed (39 out of 46); six participants were left-handed, and one had bilateral handedness. The majority of participants (32 out of 46) were left language-dominant, as determined by Wada test. Two participants were right hemisphere-dominant, and one had bilateral language dominance. The remaining 11 participants were not evaluated for language dominance; 9 of them were right-handed and thus were assumed left language-dominant for the purposes of the analysis of lateralization (see below). The remaining two participants who did not undergo Wada test and who were left-handed were not included in that analysis.

All participants underwent audiological and neuropsychological assessment prior to electrode implantation, and none had auditory or cognitive deficits that would impact the results of this study. The participants were tapered off their antiepileptic drugs during chronic monitoring when RS data were collected.

Experimental procedures

Pre-implantation neuroimaging. All participants underwent whole-brain high-resolution T1-weighted structural MRI scans before electrode implantation. In a subset of ten participants (Supplementary Table 2), RS-fMRI data were used for estimates of functional connectivity. The scanner was a 3T GE Discovery MR750W with a 32-channel head coil. The pre-electrode implantation anatomical T1 scan (3D FSPGR BRAVO sequence) was obtained with the following parameters: FOV = 25.6 cm, flip angle = 12 deg., TR = 8.50 ms, TE = 3.29 ms, inversion time = 450 ms, voxel size = 1.0 × 1.0 × 0.8 mm. For RS-fMRI, 5 blocks of 5-minute gradient-echo EPI runs (650 volumes) were collected with the following parameters: FOV = 22.0 cm, TR = 2260 ms, TE = 30 ms, flip angle = 80 deg., voxel size = 3.45 × 3.45 × 4.0 mm. In some cases, fewer RS acquisition sequences were used in the final analysis due to movement artifact or because the full scanning session was not completed. For each participant, RS-fMRI runs were acquired in the same session but non-contiguously (dispersed within an imaging session to avoid habituation). Participants were asked to keep their eyes open, and a fixation cross was presented through a projector.

513 *iEEG recordings.* iEEG recordings were obtained using either subdural and depth electrodes, or
514 depth electrodes alone, based on clinical indications. Electrode arrays were manufactured by
515 Ad-Tech Medical (Racine, WI). Subdural arrays, implanted in 36 participants out of 46, consisted
516 of platinum-iridium discs (2.3 mm diameter, 5-10 mm inter-electrode distance), embedded in a
517 silicon membrane. Stereotactically implanted depth arrays included between 4 and 12
518 cylindrical contacts along the electrode shaft, with 5-10 mm inter-electrode distance. A
519 subgaleal electrode, placed over the cranial vertex near midline, was used as a reference in all
520 participants. All electrodes were placed solely on the basis of clinical requirements, as
521 determined by the team of epileptologists and neurosurgeons[99].

522 No-task RS data were recorded in the dedicated, electrically shielded suite in The University of
523 Iowa Clinical Research Unit while the participants lay in the hospital bed. RS data were collected
524 6.4 +/- 3.5 days (mean ± standard deviation; range 1.5 – 20.9) after electrode implantation
525 surgery. In the first 15 participants (L275 through L362), data were recorded using a TDT RZ2
526 real-time processor (Tucker-Davis Technologies, Alachua, FL). In the remaining 31 participants
527 (R369 through L585), data acquisition was performed using a Neuralynx Atlas System
528 (Neuralynx Inc., Bozeman, MT). Recorded data were amplified, filtered (0.1–500 Hz bandpass, 5
529 dB/octave rolloff for TDT-recorded data; 0.7–800 Hz bandpass, 12 dB/octave rolloff for
530 Neuralynx-recorded data) and digitized at a sampling rate of 2034.5 Hz (TDT) or 2000 Hz
531 (Neuralynx). The durations of recordings were 13 +/- 11 min. In all but two participants,
532 recording durations were between 10 and 22 min.; in one participant duration was 6 min., and
533 in one participant the duration was 81 min.

534

535 *Data analysis*

536 *Anatomical reconstruction and ROI parcellation.* Localization of recording sites and their
537 assignment to ROIs relied on post-implantation T1-weighted anatomical MRI and post-
538 implantation computed tomography (CT). All images were initially aligned with pre-operative T1
539 scans using linear coregistration implemented in FSL (FLIRT)[100]. Electrodes were identified in the
540 post-implantation MRI as magnetic susceptibility artifacts and in the CT as metallic
541 hyperdensities. Electrode locations were further refined within the space of the pre-operative
542 MRI using three-dimensional non-linear thin-plate spline warping[101], which corrected for post-
543 operative brain shift and distortion. The warping was constrained with 50-100 control points,
544 manually selected throughout the brain, which were visually aligned to landmarks in the pre-
545 and post-implantation MRI.

546 To pool data across participants, the dimensionality of connectivity matrices was reduced by
547 assigning electrodes to one of 58 ROIs organized into 6 ROI groups (see Fig. 1; Supplementary
548 Table 2, 3) based upon anatomical reconstructions of electrode locations in each participant.
549 For subdural arrays, ROI assignment was informed by automated parcellation of cortical
550 gyri[102,103] as implemented in the FreeSurfer software package. For depth arrays, it was informed

551 by MRI sections along sagittal, coronal, and axial planes. For recording sites in Heschl's gyrus,
552 delineation of the border between core auditory cortex adjacent non-core areas (HGPM and
553 HGAL, respectively) was performed in each participant using physiological criteria[104,105].
554 Specifically, recording sites were assigned to HGPM if they exhibited phase-locked (frequency-
555 following) responses to 100 Hz click trains and if the averaged evoked potentials to these
556 stimuli featured short-latency (<20 ms) peaks. Such response features are characteristic for
557 HGPM and are not present within HGAL[104]. Additionally, correlation coefficients between
558 average evoked potential waveforms recorded from adjacent sites were examined to identify
559 discontinuities in response profiles along Heschl's gyrus that could be interpreted as reflecting a
560 transition from HGPM to HGAL. Superior temporal gyrus was subdivided into posterior and
561 middle non-core auditory cortex ROIs (STGP and STGM), and auditory-related anterior ROI
562 (STGA) using the transverse temporal sulcus and ascending ramus of the Sylvian fissure as
563 macroanatomical boundaries. The insula was subdivided into posterior and anterior ROIs, with
564 the former considered within the auditory-related ROI group[5]. Middle and inferior temporal
565 gyrus were each divided into posterior, middle, and anterior ROIs by diving the gyrus into three
566 approximately equal-length thirds. Angular gyrus was divided into posterior and anterior ROIs
567 using the angular sulcus as a macroanatomical boundary. Anterior cingulate cortex was
568 identified by automatic parcellation in FreeSurfer and was considered as part of the prefrontal
569 ROI group, separately from the rest of the cingulate gyrus. Postcentral and precentral gyri were
570 each divided into ventral and dorsal portions using the y_{MNI} coordinate (see below) of 40 mm as
571 a boundary. Recording sites identified as seizure foci or characterized by excessive noise, and
572 depth electrode contacts localized to the white matter or outside brain, were excluded from
573 analyses and are not listed in Supplementary Table 2.

574 *Preprocessing of fMRI data.* Standard preprocessing was applied to the RS-fMRI data acquired
575 in the pre-implantation scan using FSL's FEAT pipeline, including spatial alignment and nuisance
576 regression. White matter, cerebrospinal fluid and global ROIs were created using deep white
577 matter, lateral ventricles and a whole brain mask, respectively. Regression was performed using
578 the time series of these three nuisance ROIs as well as 6 motion parameters (3 rotations and 3
579 translations) and their derivatives, detrended with second order polynomials. Temporal
580 bandpass filtering was 0.008–0.08 Hz. Spatial smoothing was applied with a Gaussian kernel (6
581 mm full-width at half maximum). The first two images from each run were discarded. Frame
582 censoring was applied when the Euclidean norm of derivatives of motion parameters exceeded
583 0.5 mm[106]. All runs were processed in native EPI space, then the residual data were
584 transformed to MNI152 and concatenated.

585 *Preprocessing of iEEG data*. Analysis of iEEG data was performed using custom software written
586 in MATLAB Version 2020a programming environment (MathWorks, Natick, MA, USA). After
587 initial rejection of recording sites identified as seizure foci, several automated steps were taken
588 to exclude recording channels and time intervals contaminated by noise. First, channels were
589 excluded if average power in any frequency band (broadband, delta, theta, alpha, beta, gamma,
590 or high gamma; see below) exceeded 3.5 standard deviations of the average power across all

channels for that participant. Next, transient artifacts were detected by identifying voltage deflections exceeding 10 standard deviations on a given channel. A time window was identified extending before and after the detected artifact until the voltage returned to the zero-mean baseline plus an additional 100 ms buffer before and after. High-frequency artifacts were also removed by masking segments of data with high gamma power exceeding 5 standard deviations of the mean across all segments. Only time bins free of these artifact masks were considered in subsequent analyses. Artifact rejection was applied across all channels simultaneously so that all connectivity measures were derived from the same time windows. Occasionally, particular channels survived the initial average power criteria yet had frequent artifacts that led to loss of data across all the other channels. There is a tradeoff in rejecting artifacts (losing time across all channels) and rejecting channels (losing all data for that channel). If artifacts occur on many channels, there is little benefit to excluding any one channel. However, if frequent artifacts occur on one or simultaneously on up to a few channels, omitting these can save more data from other channels than those channels contribute at all other times. We chose to optimize the total data retained, channels × time windows, and omitted some channels when necessary. To remove shared signals unlikely to derive from brain activity, data from retained channels were high-pass filtered above 200 Hz, and a spatial filter was derived from the singular value decomposition omitting the first singular vector. This spatial filter was then applied to the broadband signal to remove this common signal.

Connectivity analysis. For RS-fMRI data, BOLD signals were averaged across voxel groupings and functional connectivity was calculated as Pearson correlation coefficients. Voxel groupings were either based on the Schaefer-Yeo 400 parcellation scheme[53] in MNI-152 space, or were based on iEEG electrode location in participant space (see Fig. 1). For the latter, fMRI voxels were chosen to represent comparable regions of the brain recorded by iEEG electrodes. For each electrode, the anatomical coordinates of the recording site were mapped to the closest valid MRI voxel, *E*, and a sphere of 25 voxels (25 mm^3) centered on *E* used as the corresponding recording site. This process was repeated for all *N* electrodes in the same ROI, and a single time series computed as the average of the fMRI BOLD signal in these *N*×25 voxels. These averages were used to compute an ROI-by-ROI connectivity matrix for RS-fMRI data. For comparisons between iEEG and fMRI embeddings, voxels were processed in participant space and ROI labels from the parcellation scheme illustrated in Figure 1 and Supplementary Table 2 were applied to the fMRI data. For comparisons between fMRI embeddings derived from all cortical ROIs versus fMRI embeddings derived from just ROIs sampled in the iEEG experiments, electrode locations were transformed from participant space to MNI-152 space, then assigned to ROIs within the Schaefer-Yeo 400 scheme.

For iEEG data, two measures of functional (non-directed) connectivity were used: the orthogonalized band power envelope correlation[29] and the debiased weighted phase lag index (wPLI;[52]), a measure of phase synchronization. Both measures avoid artifacts due to volume conduction by discounting connectivity near zero phase lag. For both measures, data were

divided into 60-second segments, pairwise connectivity estimated in each segment, and then connectivity estimates averaged across all segments for that participant.

Envelope correlations were estimated for each data segment and every recording site as in[29], except time-frequency decomposition was performed using the demodulated band transform[107], rather than wavelets. For each frequency band (theta: 4-8 Hz, alpha: 8-13 Hz, beta: 13-30 Hz, gamma: 30-70 Hz; high gamma: 70-120 Hz), the power at each time bin was calculated as the average (across frequencies) log of the squared amplitude. For each pair of signals *X* and *Y*, one was orthogonalized to the other by taking the magnitude of the imaginary component of the product of one signal with the normalized complex conjugate of the other:

$$Y_{orth} = |\mathrm{Im}\{Y \times X^{*}/|X|\}|$$

Both signals were band-pass filtered (0.2 – 1 Hz), and the Pearson correlation calculated between signals. The process was repeated by orthogonalizing in the other direction and the overall envelope correlation for a pair of recording sites was the average of the two Pearson correlations.

wPLI was estimated for each data segment and every recording site pair from the sign of the imaginary part of the cross-spectrum at each frequency and averaged across frequencies within each band of interest (theta: 4-8 Hz, alpha: 8-13 Hz, beta: 13-30 Hz). The cross spectrum was calculated from the demodulated band transform as described previously[108].

Prior to diffusion map embedding, connectivity matrices were thresholded by saving at least the top third (rounded up) connections for every row, as well as their corresponding columns (to preserve symmetry). We also included any connections making up the minimum spanning tree of the graph represented by the elementwise reciprocal of the connectivity matrix to ensure the graph is connected.

ROI-based connectivity analysis. Connectivity between ROIs was computed as the average envelope correlation or wPLI value between all pairs of recording sites in the two ROIs. For analyses in which connectivity was summarized across participants (Fig. 3-8), we used only a subset of ROIs such that every possible pair of included ROIs was represented in at least two participants (Supplementary Table 2). This list of ROIs was obtained by iteratively removing ROIs with the worst cross-coverage with other ROIs until every ROI remaining had sufficient coverage with all remaining ROIs.

Diffusion map embedding. See the Appendix for details about DME.

In brief, the connectivity matrix **K =** [$k(i,j)$] (here orthogonalized power envelope correlations or wPLI) is normalized by degree to yield a matrix $\mathbf{P} = \mathbf{D}^{-1}\mathbf{K}$, where **D** is the degree matrix, i.e. the diagonal elements of $\mathbf{D} = \sum_{j=1}^{N} k(i,j)$, where *N* is the number of recording sites, and the off-diagonal elements of **D** are zero. If the recording sites are conceptualized as nodes on a graph with edges defined by **K**, then **P** can be understood as the transition probability matrix for a

'random walk' or a 'diffusion' on the graph (see Appendix; [30,31]). DME consists of mapping the recording sites into an embedding space using an eigendecomposition of **P**,

$$\Psi^{(t)}(x_i) = [\lambda_1{}^t\psi_1(x_i), \lambda_1{}^t\psi_1(x_i), \ldots, \lambda_M{}^t\psi_M(x_i)]^T ,$$

where ψ_j are the eigenvectors of **P**. The parameter t is the number of time steps in that random walk; here, larger values of t correspond to exploring the structure of the data at larger spatial scales. In the analyses presented here, $k(i,j)$ is either the orthogonalized power envelope correlations, or the debiased wPLI. We note that in recent applications of DME to fMRI functional connectivity data, **K** was first transformed by applying cosine similarity[32]. However, this additional step has not been universally implemented (e.g.,[109]), nor is it required within the mathematical framework of DME[30,31]. In the case of the dataset presented here, applying cosine similarity to the data served to smear the data in embedding space, increasing variance, and decreasing separation between ROIs, but otherwise produced qualitatively similar results.

DME can be implemented alternatively based on a symmetric version of diffusion matrix $\mathbf{P_{symm}} = \mathbf{D}^{-0.5}\mathbf{K}\mathbf{D}^{-0.5}$. Basing DME on $\mathbf{P_{symm}}$ has the advantage that the eigenvectors of $\mathbf{P_{symm}}$ form an orthogonal basis set (unlike the eigenvectors of **P**), providing some additional convenience mathematically that is beyond the scope of this paper [31]. Additionally, the eigenvalues of **P** and $\mathbf{P_{symm}}$ are identical.

In two sets of analyses presented here, pairs of embeddings were compared to each other: in the analysis of lateralization of speech and language networks, and in the comparison between iEEG and fMRI data. To do that, we used a change of basis operator to map embeddings into a common embedding space using the method described in Coifman et al 2014[31].

Dimensionality reduction via low rank approximations to $\mathbf{P_{symm}}$. Diffusion map embedding offers an opportunity to reduce the dimensionality of the underlying data by considering only those dimensions that contribute importantly to the structure of the data, as manifested in the structure of the transition probability matrix **P**, or, equivalently, of the diffusion matrix $\mathbf{P_{symm}}$. We used the eigenvalue spectrum of $\mathbf{P_{symm}}$ to determine its ideal low rank approximation, balancing dimensionality reduction and information loss. The basis for this is most easily understood in terms of the eigenvalue spectrum of **P**, whose spectrum is identical to that of $\mathbf{P_{symm}}$[31]. Because **P** is real and symmetric, the magnitude of the eigenvalues is identical to the singular values of **P**. The singular values tell us about the fidelity of low rank approximations to **P**. Specifically, if **P** has a set of singular values $\sigma_1 \geq \sigma_1 \geq \ldots \geq \sigma_n$, then for any integer $k \geq 1$,

$$\min_{\widetilde{\mathbf{P}_k}} \left\| \mathbf{P} - \widetilde{\mathbf{P}_k} \right\|_2 = \sigma_{k+1},$$

where $\widetilde{\mathbf{P}_k}$ is the rank-k approximation to **P**. Thus, the magnitude of the eigenvalues corresponds to the fidelity of the lower dimensional approximation, and the difference in the magnitude of successive eigenvalues represents the improvement in that approximation as the dimensionality increases. The spectrum of **P** invariably has an inflection point ("elbow"), separating two sets of eigenvalues λ_i: those whose magnitude decreases more quickly with

increasing i, and those beyond the inflection point whose magnitude decreases more slowly with increasing i. The inflection point thus delineates the number of dimensions that are most important for approximating $\mathbf{P}$ or $\mathbf{P}_{\mathbf{symm}}$. The inflection point k_{infl} was identified algorithmically[110], and the number of dimensions retained set equal to $k_{\text{infl}} - 1$.

Comparing distances in embedding space. The relative distance between points in embedding space provides insight into the underlying functional geometry. In several analyses presented here, two embeddings of identical sets of ROIs were compared as ROI distances within the two embeddings. After mapping to a common space and reducing dimensionality as described above, the two embeddings A and B were used to create the pairwise distance matrices A\` and B\`. The Pearson correlation coefficient r was then computed between the upper triangles (excluding the diagonal) of the corresponding elements in the distance matrices.

Signal to noise (SNR) characteristics. To measure the robustness of the embedding analysis to variability over time, an SNR was computed as follows. For each participant, a channel × channel $\mathbf{P}_{\mathbf{symm}}$ matrix was calculated for each 60 s segment of data. For each segment, DME analysis was applied and a channel × channel distance matrix calculated. These distance matrices were averaged across segments. The 'signal' of interest was defined as the variability (standard deviation) of this averaged distance matrix (ignoring the diagonals). The 'noise' was defined as the variability across time, estimated for each element of the distance matrix as the standard deviation across segments, then averaged across the elements of the matrix. The SNR for functional connectivity itself was computed in an analogous manner, using the original channel × channel connectivity matrix rather than the matrix of embedding distances.

Estimating precision in position and distances in embedding space. To obtain error estimates for both ROI locations in embedding space and embedding distance between ROIs, average ROI × ROI adjacency matrices were calculated. This was done by drawing each edge from an averaged bootstrap sample across participants, obtaining 10,000 such adjacency matrices, and performing diffusion map embedding for each. For locations in embedding space, these embeddings were then mapped via the change of basis procedure described above to the original group average embedding space. For each ROI, the mapped bootstrap iterations produced a cloud of locations in embedding space which were summarized by the standard deviation in each dimension. For embedding distances, no change of basis was necessary because distances were preserved across bases.

To compare the positions of STSL versus STSU relative to canonical auditory cortical ROIs (HGPM, HGAL, PT, PP, STGP, and STGM) or ROIs involved in semantic processing (STGA, MTGA, MTGP, ITGA, ITGP, TP, AGA, AGP, SMG, IFGop, IFGtr, IFGor[28,38-40]), we calculated the average pairwise distance from STSL or STSU to each such ROI. The difference between these averages was compared to a null distribution obtained by Monte Carlo sampling of the equivalent statistic obtained by randomly exchanging STSL/STSU labels by participant. The specific comparisons performed were chosen *a priori* to constrain the number of possible hypotheses to test; pairwise comparisons of all possible ROI pairs (let alone comparisons of all higher-order

742 groupings) would not have had sufficient statistical power under appropriate corrections for
743 multiple comparisons. Though different choices could have been made for inclusion in the
744 "semantic processing" category, exchanging one or two of these ROIs would not strongly
745 influence the average distance in a group of twelve ROIs.

746 *Hierarchical clustering.* Agglomerative hierarchical clustering was done using the *linkage*
747 function in MATLAB, with Euclidean distance as the distance metric and Ward's linkage
748 (minimum variance algorithm) as the linkage method. The ordering of ROIs along the horizontal
749 axis in the dendrogram was determined using the *optimalleaforder* function in MATLAB, with
750 the optimization criterion set to 'group'.

751 *Comparing language dominant/non-dominant hemispheres.* To test for differences in
752 embedding location between language dominant and non-dominant hemispheres, two
753 measures were considered: different location of individual ROIs in embedding space, and
754 different pairwise distances between ROIs in embedding space. To calculate differences in
755 location of individual ROIs, dominant/non-dominant average embeddings were mapped to a
756 common space (from an embedding using the average across all participants regardless of
757 language dominance) using the change of basis operator. The language-dominant location
758 difference for a specific ROI was calculated as the Euclidean distance between the two locations
759 of each ROI in this common space. For pairwise distances, the change of basis is irrelevant, so
760 pairwise Euclidean distances were calculated in embedding space for each hemisphere and
761 then subtracted to obtain a difference matrix. To determine whether the differences in location
762 or pairwise distances were larger than expected by chance, random permutations of the
763 dominant/non-dominant labels were used to generate empirical null distributions. Since this
764 approach produces a *p*-value for every pair of connections, *p*-values were adjusted using false
765 discovery rate (FDR) to account for multiple comparisons.

766 *Analyses of fMRI connectivity in embedding space.* Two sets of analyses were performed using
767 fMRI data. First, iEEG and fMRI data were compared in embedding space. In this analysis,
768 connectivity based on RS-fMRI data from voxels located at electrode recording sites was
769 compare with the corresponding connectivity matrix derived from iEEG data. The embedding
770 analysis was applied to the two connectivity matrices, all pairwise inter-ROI distances
771 computed, and iEEG and fMRI data compared using the correlation of the pairwise ROI
772 distances. The second analysis was to compare embeddings derived from all ROIs in the RS-
773 fMRI scans to those derived from just ROIs sampled with iEEG electrodes. Here, ROI × ROI
774 connectivity matrices were computed for all ROIs, then embeddings created from the full
775 matrices or from matrices containing just rows and columns corresponding to the ROIs sampled
776 with iEEG.

Data and code availability

Data is available via a request to the Authors pending establishment of a formal data sharing agreement, submission of a formal project outline, and agreement about co-authorship. All software developed for this study will be freely available via Git repository. Please contact Bryan Krause (bmkrause@wisc.edu) for details.

Appendix: Diffusion Map Embedding

In the framework of DME, we consider a space X that is the set of N recording sites. We compute the similarity between those sites based on the time varying signals recorded at each site, defining similarity as the functional connectivity measure $k(x_i,x_j)$.

Define the matrix **K** whose i,j^{th} element is $k(x_i,x_j)$. $k(x_i,x_j)$ is required to be symmetric, i.e., $k(x_i,x_j) = k(x_j,x_i)$, and positivity preserving, i.e. $k(x_i,x_j) > 0$ for all $[i,j]$, to allow for spectral analysis of a normalized version of **K**.

From X and **K** we can construct a weighted graph Γ in which the vertices are the nodes and the edge weights are $k(x_i,x_j)$. We take random walks on the graph at time steps $t = 1, 2, \ldots$, jumping from node x_i to node x_j at each time step, with the (stochastic) decision as to which node should be visited next depending on $k(x_i,x_j)$.

Define

$$p(x_i,x_j) = k(x_i,x_j)/d(x_i),$$

where

$$d(x_i) = \Sigma_j[k(x_i,x_j)]$$

is the degree of node x_i. Normalizing $k(x_i,x_j)$ in this way allows us to interpret it as the probability $p(x_i,x_j)$ that we'll jump from vertex x_i to vertex x_j in a single time step of our random walk.

If we consider a single time step, we only capture the structure in X on a very local scale, since we can only jump between vertices that are directly connected. As we run the random walk forward in time, we begin to explore more of our neighborhood, and we begin to explore other neighborhoods as well. Two vertices x_i and x_j that have similar connectivity to the rest of the network have a high probability of being connected during these longer walks because they themselves are connected to similar groups of vertices, and so there are many possible paths between x_i and x_j.

The diffusion operator (matrix) $\mathbf{P} = [p(x_i,x_j)]$ describes how signals diffuse from node to node in the graph. If **v** is a $N\times1$ vector (i.e., a value assigned to each vertex, for example representing an input to each node), then **P** describes what will happen to that input as time goes on.

$$\mathbf{Pv} = [p(x_1,x_1)\mathbf{v}[x_1]+p(x_1,x_2)\mathbf{v}[x_2]+\ldots;\ p(x_2,x_1)\mathbf{v}[x_1]+p(x_2,x_2)\mathbf{v}[x_2]+\ldots;\ldots]^T$$

If, for example, all the nodes were insular, with $p(x_i,x_i)=1$ for all i, and otherwise $p(x_i,x_j)=0$, $\mathbf{Pv} = \mathbf{v}$, i.e., no diffusion occurs. If the probabilities are more distributed, **Pv** would reveal how much signals diffuse out from each node given the starting condition of **v**. Importantly, $\mathbf{P}^k\mathbf{v}$ would reveal what that distribution looks like after k time steps.

The eigenvector expansion of **P** based on its eigenvectors ψ_j and eigenvalues , λ_j, j = 1...N, is a natural method for uncovering structure in **P** because each eigenvector of **P** is a dimension along which relevant organizational features emerge. That is, clusters of related points (communities) tend to be distinct and ordered along these dimensions. In fact, we could preserve a lot of information about **P** by keeping just a subset of M of these vectors and discarding the rest. The information we want to preserve in the context of diffusion map embedding is the functional distance between the data at two nodes given t time steps to meander through the graph. We can define the diffusion map

$$\Psi^{(t)}(x_i) = [\lambda_1{}^t\psi_1(x_i), \lambda_1{}^t\psi_1(x_i), \ldots, \lambda_M{}^t\psi_M(x_i)]^T,$$

which maps each point x in X to a point in an embedding space of dimension $M \leq$ N. In this space, the diffusion distance D, which is the Euclidean distance between points, is the difference in the probability distributions linking x_i to the rest of the network and x_j to the rest of the network:

$$D^{(t)}(x_i,x_j)^2 = ||\Psi^{(t)}(x_i) - \Psi^{(t)}(x_j)||_{l2}.= ||p^{(t)}(x_i,:)- p^{(t)}(x_j,:)||^2{}_{l2}.$$

We return now to the parameter t, which corresponds to the time scale of the diffusion process (i.e., the number of steps in the random walk on the graph). As t progresses, the coordinates of the data in embedding space are scaled according to $\lambda_i{}^t$, where λ_i is the eigenvalue of the ith dimension being scaled. Thus, the value of t sets the spatial scale of the analysis, with higher values de-emphasizing smaller eigenvalues. Because $|\lambda_i|<1\ \forall\ i$, at higher values of t each dimension will be scaled down ('collapse'), with the dimension corresponding to max($|\lambda_i|$) (i.e., λ_1) scaled the least.

To compare embeddings across groups of participants, or modalities of measurements, it is necessary to map embeddings to a common space. To do so, consider two sets of data α and β, and the data spaces X_α and X_β. The problem is that X_α and X_β are different spaces with different kernels k_α and k_β. This means that the eigenvectors for $\mathbf{P}_\alpha$ and $\mathbf{P}_\beta$ will be different, and data projected into a space defined by some subset of the eigenvectors cannot be compared directly. The solution is to apply a change of basis operator to one set of the eigenvectors to get the data into the same embedding space[31]:

$$D^{(t)}(x_{i|\alpha},x_{i|\beta}) = ||\Psi^{(t)}{}_\alpha(x) - O_{\beta\rightarrow\alpha}\Psi^{(t)}{}_\beta(x)||_{l2}.$$

Where the change of basis operator $O_{\beta\rightarrow\alpha}$ is defined as

$$O_{\beta\rightarrow\alpha}\ \mathbf{v} = \Sigma_j[\mathbf{v}(j)<\psi^{(i)}{}_\alpha\ ,\psi^{(j)}{}_\beta>]_{i>=1},$$

Where <**a**,**b**> is the inner product of **a** and **b**.

REFERENCES CITED

1. Scott, S.K. The neurobiology of speech perception and production--can functional imaging tell us anything we did not already know? *J Commun Disord* **45**, 419-425 (2012).
2. Woods, D.L. & Alain, C. Functional imaging of human auditory cortex. *Curr Opin Otolaryngol Head Neck Surg* **17**, 407-411 (2009).
3. Friederici, A.D., Meyer, M. & von Cramon, D.Y. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang* **75**, 289-300 (2000).
4. Angulo-Perkins, A.*, et al.* Music listening engages specific cortical regions within the temporal lobes: differences between musicians and non-musicians. *Cortex* **59**, 126-137 (2014).
5. Zhang, Y.*, et al.* The Roles of Subdivisions of Human Insula in Emotion Perception and Auditory Processing. *Cereb Cortex* **29**, 517-528 (2019).
6. Abrams, D.A., Kochalka, J., Bhide, S., Ryali, S. & Menon, V. Intrinsic functional architecture of the human speech processing network. *Cortex* **129**, 41-56 (2020).
7. Nourski, K.V.*, et al.* Electrophysiology of the Human Superior Temporal Sulcus during Speech Processing. *Cereb Cortex* **31**, 1131-1148 (2021).
8. Rauschecker, J.P. & Scott, S.K. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* **12**, 718-724 (2009).
9. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat Rev Neurosci* **8**, 393-402 (2007).
10. Friederici, A.D. The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn Sci* **16**, 262-268 (2012).
11. Cloutman, L.L. Interaction between dorsal and ventral processing streams: where, when and how? *Brain Lang* **127**, 251-263 (2013).
12. Hickok, G. & Poeppel, D. Neural basis of speech perception. *Handb Clin Neurol* **129**, 149-160 (2015).
13. Rauschecker, J.P. Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex* **98**, 262-268 (2018).
14. Munoz-Lopez, M.M., Mohedano-Moriano, A. & Insausti, R. Anatomical pathways for auditory memory in primates. *Front Neuroanat* **4**, 129 (2010).
15. Kraus, K.S. & Canlon, B. Neuronal connectivity and interactions between the auditory and limbic systems. Effects of noise and tinnitus. *Hear Res* **288**, 34-46 (2012).
16. Husain, F.T. & Schmidt, S.A. Using resting state functional connectivity to unravel networks of tinnitus. *Hear Res* **307**, 153-162 (2014).
17. Kumar, S.*, et al.* A Brain System for Auditory Working Memory. *J Neurosci* **36**, 4492-4505 (2016).
18. Kumar, S.*, et al.* Oscillatory correlates of auditory working memory examined with human electrocorticography. *Neuropsychologia* **150**, 107691 (2021).
19. Geschwind, N. The organization of language and the brain. *Science* **170**, 940-944 (1970).
20. Hagoort, P. The neurobiology of language beyond single-word processing. *Science* **366**, 55-58 (2019).
21. McGettigan, C. & Scott, S.K. Cortical asymmetries in speech perception: what's wrong, what's right and what's left? *Trends Cogn Sci* **16**, 269-276 (2012).
22. Turkeltaub, P.E. & Coslett, H.B. Localization of sublexical speech perception components. *Brain Lang* **114**, 1-15 (2010).
23. Leaver, A.M. & Rauschecker, J.P. Cortical representation of natural complex sounds: effects of acoustic features and auditory object category. *J Neurosci* **30**, 7604-7612 (2010).

24. Eisner, F., McGettigan, C., Faulkner, A., Rosen, S. & Scott, S.K. Inferior frontal gyrus activation predicts individual differences in perceptual learning of cochlear-implant simulations. *J Neurosci* **30**, 7179-7186 (2010).
25. Saur, D.*, et al.* Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A* **105**, 18035-18040 (2008).
26. Jackson, R.L., Bajada, C.J., Rice, G.E., Cloutman, L.L. & Lambon Ralph, M.A. An emergent functional parcellation of the temporal cortex. *Neuroimage* **170**, 385-399 (2018).
27. Visser, M., Jefferies, E. & Lambon Ralph, M.A. Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci* **22**, 1083-1094 (2010).
28. Ralph, M.A., Jefferies, E., Patterson, K. & Rogers, T.T. The neural and computational bases of semantic cognition. *Nat Rev Neurosci* **18**, 42-55 (2017).
29. Hipp, J.F., Hawellek, D.J., Corbetta, M., Siegel, M. & Engel, A.K. Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nat Neurosci* **15**, 884-890 (2012).
30. Coifman, R.R.*, et al.* Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. *Proc Natl Acad Sci U S A* **102**, 7426-7431 (2005).
31. Coifman, R.R. & Hirn, M.J. Diffusion maps for changing data. *Applied and Computational Harmonic Analysis* **36**, 79-107 (2014).
32. Margulies, D.S.*, et al.* Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci U S A* **113**, 12574-12579 (2016).
33. Jones, E.G., Coulter, J.D. & Hendry, S.H. Intracortical connectivity of architectonic fields in the somatic sensory, motor and parietal cortex of monkeys. *J Comp Neurol* **181**, 291-347 (1978).
34. Morel, A., Garraghty, P.E. & Kaas, J.H. Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J Comp Neurol* **335**, 437-459 (1993).
35. Kaas, J.H. & Hackett, T.A. Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci USA* **97**, 11793-11799 (2000).
36. Kaas, J.H. & Hackett, T.A. Subdivisions of auditory cortex and levels of processing in primates. *Audiol Neurootol* **3**, 73-85 (1998).
37. Cavada, C., Company, T., Tejedor, J., Cruz-Rizzolo, R.J. & Reinoso-Suarez, F. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb Cortex* **10**, 220-242 (2000).
38. Binder, J.R., Desai, R.H., Graves, W.W. & Conant, L.L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* **19**, 2767-2796 (2009).
39. Humphreys, G.F., Hoffman, P., Visser, M., Binney, R.J. & Lambon Ralph, M.A. Establishing task- and modality-dependent dissociations between the semantic and default mode networks. *Proc Natl Acad Sci U S A* **112**, 7857-7862 (2015).
40. Jackson, R.L., Hoffman, P., Pobric, G. & Lambon Ralph, M.A. The Semantic Network at Work and Rest: Differential Connectivity of Anterior Temporal Lobe Subregions. *J Neurosci* **36**, 1490-1501 (2016).
41. Remedios, R., Logothetis, N.K. & Kayser, C. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. *J Neurosci* **29**, 1034-1045 (2009).
42. Steinschneider, M., Nourski, K.V. & Fishman, Y.I. Representation of speech in human auditory cortex: is it special? *Hear Res* **305**, 57-73 (2013).
43. Craig, A.D. Interoception: the sense of the physiological condition of the body. *Curr Opin Neurobiol* **13**, 500-505 (2003).
44. Kuehn, E., Mueller, K., Lohmann, G. & Schuetz-Bosbach, S. Interoceptive awareness changes the posterior insula functional connectivity profile. *Brain Struct Funct* **221**, 1555-1571 (2016).
45. Bernstein, L.E. & Liebenthal, E. Neural pathways for visual speech perception. *Front Neurosci* **8**, 386 (2014).
46. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat.Rev.Neurosci.* **10**, 186-198 (2009).

941 47. Knecht, S.*, et al.* Handedness and hemispheric language dominance in healthy humans. *Brain*
942 **123 Pt 12**, 2512-2518 (2000).
943 48. Price, C.J. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech,
944 spoken language and reading. *Neuroimage* **62**, 816-847 (2012).
945 49. Schirmer, A., Fox, P.M. & Grandjean, D. On the spatial organization of sound processing in the
946 human temporal lobe: a meta-analysis. *Neuroimage* **63**, 137-147 (2012).
947 50. Chang, E.F., Raygor, K.P. & Berger, M.S. Contemporary model of language organization: an
948 overview for neurosurgeons. *J Neurosurg* **122**, 250-261 (2015).
949 51. Ardila, A., Bernal, B. & Rosselli, M. How Localized are Language Brain Areas? A Review of
950 Brodmann Areas Involvement in Oral Language. *Arch Clin Neuropsychol* **31**, 112-122 (2016).
951 52. Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F. & Pennartz, C.M. An improved index
952 of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and
953 sample-size bias. *Neuroimage* **55**, 1548-1565 (2011).
954 53. Schaefer, A.*, et al.* Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic
955 Functional Connectivity MRI. *Cereb Cortex* **28**, 3095-3114 (2018).
956 54. Hackett, T.A., Preuss, T.M. & Kaas, J.H. Architectonic identification of the core region in auditory
957 cortex of macaques, chimpanzees, and humans. *J Comp Neurol.* **441**, 197-222 (2001).
958 55. Hackett, T.A. Anatomic organization of the auditory cortex. *Handb Clin Neurol* **129**, 27-53 (2015).
959 56. Woods, D.L.*, et al.* Functional properties of human auditory cortical fields. *Front Syst Neurosci* **4**,
960 155 (2010).
961 57. Barton, B., Venezia, J.H., Saberi, K., Hickok, G. & Brewer, A.A. Orthogonal acoustic dimensions
962 define auditory field maps in human cortex. *Proc Natl Acad Sci U S A* **109**, 20738-20743 (2012).
963 58. Moerel, M., De Martino, F. & Formisano, E. An anatomical and functional topography of human
964 auditory cortical areas. *Front Neurosci* **8**, 225 (2014).
965 59. Howard, M.A.*, et al.* Auditory cortex on the human posterior superior temporal gyrus. *J Comp*
966 *Neurol.* **416**, 79-92 (2000).
967 60. Nourski, K.V.*, et al.* Spectral organization of the human lateral superior temporal gyrus revealed
968 by intracranial recordings. *Cereb Cortex* **24**, 340-352 (2014).
969 61. Hamilton, L.S., Oganian, Y., Hall, J. & Chang, E.F. Parallel and distributed encoding of speech
970 across human auditory cortex. *Cell* **184**, 4626-4639 e4613 (2021).
971 62. Hickok, G. The functional neuroanatomy of language. *Phys Life Rev* **6**, 121-143 (2009).
972 63. Beauchamp, M.S. The social mysteries of the superior temporal sulcus. *Trends Cogn Sci* **19**, 489-
973 490 (2015).
974 64. Venezia, J.H.*, et al.* Auditory, Visual and Audiovisual Speech Processing Streams in Superior
975 Temporal Sulcus. *Front Hum Neurosci* **11**, 174 (2017).
976 65. Zachlod, D.*, et al.* Four new cytoarchitectonic areas surrounding the primary and early auditory
977 cortex in human brains. *Cortex* **128**, 1-21 (2020).
978 66. Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P. & Pike, B. Voice-selective areas in human auditory
979 cortex. *Nature* **403**, 309-312 (2000).
980 67. Deen, B., Koldewyn, K., Kanwisher, N. & Saxe, R. Functional Organization of Social Perception
981 and Cognition in the Superior Temporal Sulcus. *Cereb Cortex* **25**, 4596-4609 (2015).
982 68. Wilson, S.M., Bautista, A. & McCarron, A. Convergence of spoken and written language
983 processing in the superior temporal sulcus. *Neuroimage* **171**, 62-74 (2018).
984 69. Kahn, I., Andrews-Hanna, J.R., Vincent, J.L., Snyder, A.Z. & Buckner, R.L. Distinct cortical anatomy
985 linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. *J*
986 *Neurophysiol* **100**, 129-139 (2008).
987 70. Wang, S.F., Ritchey, M., Libby, L.A. & Ranganath, C. Functional connectivity based parcellation of
988 the human medial temporal lobe. *Neurobiol Learn Mem* **134 Pt A**, 123-134 (2016).

71. Michelmann, S.*, et al.* Moment-by-moment tracking of naturalistic learning and its underlying hippocampo-cortical interactions. *Nat Commun* **12**, 5394 (2021).
72. Rocchi, F.*, et al.* Common fronto-temporal effective connectivity in humans and monkeys. *Neuron* **109**, 852-868 e858 (2021).
73. Fruhholz, S., Trost, W. & Kotz, S.A. The sound of emotions-Towards a unifying neural network perspective of affective sound processing. *Neurosci Biobehav Rev* **68**, 96-110 (2016).
74. Upadhyay, J.*, et al.* Effective and structural connectivity in the human auditory cortex. *J Neurosci* **28**, 3341-3349 (2008).
75. Maller, J.J.*, et al.* Revealing the Hippocampal Connectome through Super-Resolution 1150-Direction Diffusion MRI. *Sci Rep* **9**, 2418 (2019).
76. Munoz-Lopez, M., Insausti, R., Mohedano-Moriano, A., Mishkin, M. & Saunders, R.C. Anatomical pathways for auditory memory II: information from rostral superior temporal gyrus to dorsolateral temporal pole and medial temporal cortex. *Front Neurosci* **9**, 158 (2015).
77. Olson, I.R., McCoy, D., Klobusicky, E. & Ross, L.A. Social cognition and the anterior temporal lobes: a review and theoretical framework. *Soc Cogn Affect Neurosci* **8**, 123-133 (2013).
78. Mesulam, M.M. Paralimbic (mesocortical) areas. in *Principles of behavioral and cognitive neurology* 49-54 (Oxford University Press, New York, NY, 2000).
79. Chanes, L. & Barrett, L.F. Redefining the Role of Limbic Areas in Cortical Processing. *Trends Cogn Sci* **20**, 96-106 (2016).
80. Hickok, G. Computational neuroanatomy of speech production. *Nat Rev Neurosci* **13**, 135-145 (2012).
81. Simmons, W.K. & Martin, A. The anterior temporal lobes and the functional architecture of semantic memory. *J Int Neuropsychol Soc* **15**, 645-649 (2009).
82. Abel, T.J.*, et al.* Direct physiologic evidence of a heteromodal convergence region for proper naming in human left anterior temporal lobe. *J Neurosci* **35**, 1513-1520 (2015).
83. Patterson, K., Nestor, P.J. & Rogers, T.T. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci* **8**, 976-987 (2007).
84. Scott, S.K., Blank, C.C., Rosen, S. & Wise, R.J. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* **123 Pt 12**, 2400-2406 (2000).
85. Spitsyna, G., Warren, J.E., Scott, S.K., Turkheimer, F.E. & Wise, R.J. Converging language streams in the human temporal lobe. *J Neurosci* **26**, 7328-7336 (2006).
86. Gorno-Tempini, M.L.*, et al.* Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol* **55**, 335-346 (2004).
87. Hein, G. & Knight, R.T. Superior temporal sulcus--It's my area: or is it? *J Cogn Neurosci* **20**, 2125-2136 (2008).
88. Makris, N.*, et al.* Delineation of the middle longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cereb Cortex* **19**, 777-785 (2009).
89. Binney, R.J., Parker, G.J. & Lambon Ralph, M.A. Convergent connectivity and graded specialization in the rostral human temporal lobe as revealed by diffusion-weighted imaging probabilistic tractography. *J Cogn Neurosci* **24**, 1998-2014 (2012).
90. de Heer, W.A., Huth, A.G., Griffiths, T.L., Gallant, J.L. & Theunissen, F.E. The Hierarchical Cortical Organization of Human Speech Processing. *J Neurosci* **37**, 6539-6557 (2017).
91. Binder, J.R.*, et al.* Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* **10**, 512-528 (2000).
92. Cogan, G.B.*, et al.* Sensory-motor transformations for speech occur bilaterally. *Nature* **507**, 94-98 (2014).
93. Gonzalez Alam, T.*, et al.* A tale of two gradients: differences between the left and right hemispheres predict semantic cognition. *Brain Struct Funct* (2021).

1037 94. Karolis, V.R., Corbetta, M. & Thiebaut de Schotten, M. The architecture of functional
1038 lateralisation and its relationship to callosal connectivity in the human brain. *Nat Commun* **10**, 1417
1039 (2019).
1040 95. Sherman, S.M. & Guillery, R.W. Distinct functions for direct and transthalamic corticocortical
1041 connections. *J Neurophysiol* **106**, 1068-1077 (2011).
1042 96. Hu, B. Functional organization of lemniscal and nonlemniscal auditory thalamus. *Exp Br Res* **153**,
1043 543-549 (2003).
1044 97. Biswal, B.B.*, et al.* Toward discovery science of human brain function. *Proc Natl Acad Sci U S A*
1045 **107**, 4734-4739 (2010).
1046 98. Seitzman, B.A., Snyder, A.Z., Leuthardt, E.C. & Shimony, J.S. The State of Resting State Networks.
1047 *Top Magn Reson Imaging* **28**, 189-196 (2019).
1048 99. Nourski, K.V. & Howard, M.A., 3rd. Invasive recordings in the human auditory cortex. *Handb Clin*
1049 *Neurol* **129**, 225-244 (2015).
1050 100. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and
1051 accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825-841 (2002).
1052 101. Rohr, K.*, et al.* Landmark-based elastic registration using approximating thin-plate splines. *IEEE*
1053 *Trans Med Imaging* **20**, 526-534 (2001).
1054 102. Destrieux, C., Fischl, B., Dale, A. & Halgren, E. Automatic parcellation of human cortical gyri and
1055 sulci using standard anatomical nomenclature. *Neuroimage* **53**, 1-15 (2010).
1056 103. Destrieux, C.*, et al.* A practical guide for the identification of major sulcogyral structures of the
1057 human cortex. *Brain Struct Funct* **222**, 2001-2015 (2017).
1058 104. Brugge, J.F.*, et al.* Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J*
1059 *Neurophysiol* **102**, 2358-2374 (2009).
1060 105. Nourski, K.V., Steinschneider, M. & Rhone, A.E. Electrocorticographic Activation within Human
1061 Auditory Cortex during Dialog-Based Language and Cognitive Testing. *Front Hum Neurosci* **10**, 202
1062 (2016).
1063 106. Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L. & Petersen, S.E. Spurious but systematic
1064 correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59**, 2142-
1065 2154 (2012).
1066 107. Kovach, C.K. & Gander, P.E. The demodulated band transform. *J Neurosci Methods* **261**, 135-154
1067 (2016).
1068 108. Banks, M.I.*, et al.* Cortical functional connectivity indexes arousal state during sleep and
1069 anesthesia. *Neuroimage* **211**, 116627 (2020).
1070 109. Langs, G., Golland, P., Tie, Y., Rigolo, L. & Golby, A.J. Functional Geometry Alignment and
1071 Localization of Brain Areas. *Adv Neural Inf Process Syst* **1**, 1225-1233 (2010).
1072 110. Satopaa, V., Albrecht, J., Irwin, D. & Raghavan, B. Finding a "Kneedle" in a Haystack: Detecting
1073 Knee Points in System Behavior. in *2011 31st International Conference on Distributed Computing*
1074 *Systems Workshops* 166-171 (2011).

1075