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, S10-OD025025, 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.

Summary

Understanding central auditory processing critically depends on defining underlying auditory cortical networks and their relationship to the rest of the brain. We addressed these questions using resting state functional connectivity derived from human intracranial electroencephalography. Mapping recording sites into a low-dimensional space where proximity represents functional similarity revealed a hierarchical organization. At fine scale, an auditory cortical cluster excluded several higher order auditory areas and segregated maximally from prefrontal cortex. On mesoscale, a cluster of limbic structures in proximity to the auditory cortex suggested a limbic stream that parallels the classically described ventral and dorsal auditory processing streams. Global hubs were identified within anterior temporal and cingulate cortex, consistent with their respective roles in semantic and cognitive processing. On a macro scale, observed hemispheric asymmetries were not specific for speech and language networks. This approach can be applied to multivariate brain data with respect to development, behavior, and disorders.

Introduction

The meso- and macroscopic organization of human neocortex has been investigated extensively using resting state (RS) functional connectivity, primarily using functional magnetic resonance imaging (fMRI) (Biswal et al., 2010; Yeo et al., 2011). RS data are advantageous as they avoid the substantial confound of stimulus-driven correlations yet identify networks that overlap with those obtained using event-related data (Smith et al., 2009), and thus are relevant to cognitive and perceptual processing. RS fMRI has contributed greatly to our understanding of the organization of the human auditory cortical hierarchy (Jackson et al., 2018; Scott, 2012; Woods and Alain, 2009), but only a few complementary studies have been conducted using electrophysiology in humans (e.g. Ko et al., 2013; Wang et al., 2021; Zhang et al., 2021). Compared to fMRI, intracranial electroencephalography (iEEG) offers superior spatio-temporal resolution and is free of methodological problems that affect MRI in key regions such as the anterior temporal lobe (Lambon Ralph et al., 2017; Visser et al., 2010). However, variable electrode coverage in human intracranial patients and small sample sizes are challenges to generalizing results.

We overcome these limitations using a large cohort of subjects that together have coverage over most of the cerebral cortex and leverage these data to address outstanding questions about auditory networks. We address the organization of human auditory cortex at three spatial scales: fine-scale organization of regions adjacent to canonical auditory cortex, clustering of cortical regions into functional processing streams, and hemispheric (a)symmetry associated with language dominance. We present a unified analytical framework applied to resting state human iEEG data that embeds functional connectivity data into a Euclidean space in which proximity represents functional similarity. We extend the analytical approach as previously applied to RS fMRI (Margulies et al., 2016) and demonstrate methodology appropriate for hypothesis testing at each of these spatial scales.

At the fine scale, though there is broad agreement that posteromedial Heschl's gyrus (HGPM) represents core auditory cortex, functional relationships among HGPM and neighboring higher-order areas are still a matter of debate. 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 (Angulo-Perkins et al., 2014; Friederici et al., 2000). The posterior insula (InsP), on the other hand, has response properties similar to HGPM, yet is not considered a canonical auditory area (Zhang et al., 2019). The superior temporal sulcus (STS) is a critical node in speech and language networks (Abrams et al., 2020; Beauchamp, 2015; Chang et al., 2015; Hickok, 2009; Price, 2012; Venezia et al., 2017), yet its functional relationships with with other auditory areas are difficult to distinguish with neuroimaging methods. Indeed, the distinct roles of its upper and lower banks (STSU, STSL) have only been recently elucidated with iEEG (Nourski et al., 2021).

Questions remain regarding mesoscale organization as well. While the auditory hierarchy is posited to be organized along two processing streams (Friederici, 2012; Hickok and Poeppel,

2007; Rauschecker and Scott, 2009), the specific brain regions involved and the functional relationships within each stream are vigorously debated (Cloutman, 2013; Hickok and Poeppel, 2015; Rauschecker, 2018; Saur et al., 2008). Furthermore, communication between auditory cortex and hippocampus, amygdala, and anterior insula (InsA) (Munoz-Lopez et al., 2010) – areas involved in auditory working memory and processing of emotional aspects of auditory information (Husain and Schmidt, 2014; Kraus and Canlon, 2012; Kumar et al., 2021; Kumar et al., 2016) – suggests a third "limbic" auditory processing stream, complementary to the dorsal and ventral streams.

At a macroscopic scale, hemispheric lateralization of speech and language processing is a widely accepted organizational feature (Geschwind, 1970; Hagoort, 2019). However, the degree to which lateralization shapes the auditory hierarchy and is reflected in hemisphere-specific connectivity profiles is unknown(Eisner et al., 2010; Hickok and Poeppel, 2015; Leaver and Rauschecker, 2010; McGettigan and Scott, 2012; Rauschecker and Scott, 2009; Turkeltaub and Coslett, 2010).

To address these questions, we applied diffusion map embedding (DME) (Coifman and Hirn, 2014; Coifman et al., 2005) to functional connectivity measured between cortical regions of interest (ROIs). DME maps connectivity into functional geometry: relationships in a Euclidean space where proximity of two ROIs reflects similarity in connectivity to the rest of the network. The DME approach provides a low-dimensional representation convenient for display while also facilitating quantitative comparisons on multiple spatial scales, including permutation-based hypothesis testing of specific ROI relationships, hierarchical clustering to identify functional processing streams, and contrasts of whole embeddings between participant cohorts.

Results

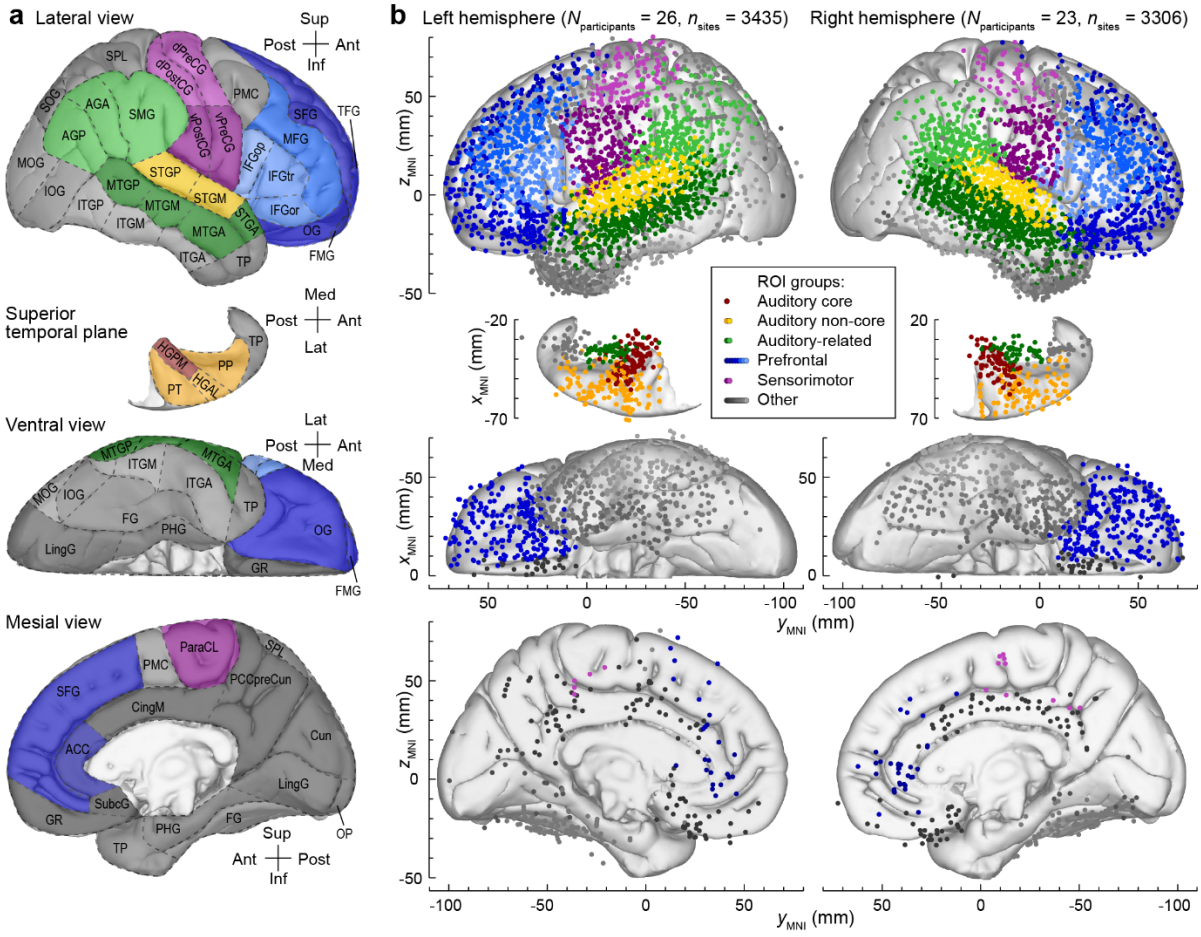
DME applied to iEEG data

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 49 participants (22 female; 6741 recording sites; Fig. 1, Supplementary Tables 1, 2). On average, each participant contributed 138±54 recording sites, representing 28±7.7 ROIs (mean ± standard deviation) (see example in Fig. 2a).

The brain parcellation scheme depicted in Figure 1A was developed based on a combination of physiological and anatomical criteria, and has been useful in our previous analyses that were largely focused on auditory processing. Below, we revisit this parcellation with a data-driven scheme.

DME was applied to pairwise functional connectivity measured as orthogonalized power envelope correlations (Hipp et al., 2012) computed between recording sites in each participant. We focus primarily on gamma-band power envelope correlations, but supplement with results from other bands for comparison. 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 using the structure of $\mathbf{P}_{\mathbf{symm}}$ and a free parameter t to map the recording sites into an embedding space. In this space, proximity between nodes represents similarity in their connectivity to the rest of the network (Fig. 2c; see Supplementary Fig. 1 for additional views). The parameter t corresponds to diffusion time: larger values of t shift focus from local towards global organization. DME exhibited superior signal-to noise characteristics compared to direct analysis of functional connectivity in 43 out of 49 participants (Supplementary Fig. 2).

Functionally distinct regions are isolated along principal dimensions 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 regions (e.g., middle temporal gyrus) had a more distributed representation within the embedding space, 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	PCCpreCun
	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 49 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* = 23), parietal operculum (*n* = 21), amygdala (*n* = 80), hippocampus (*n* = 86), putamen (*n* = 15), 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.

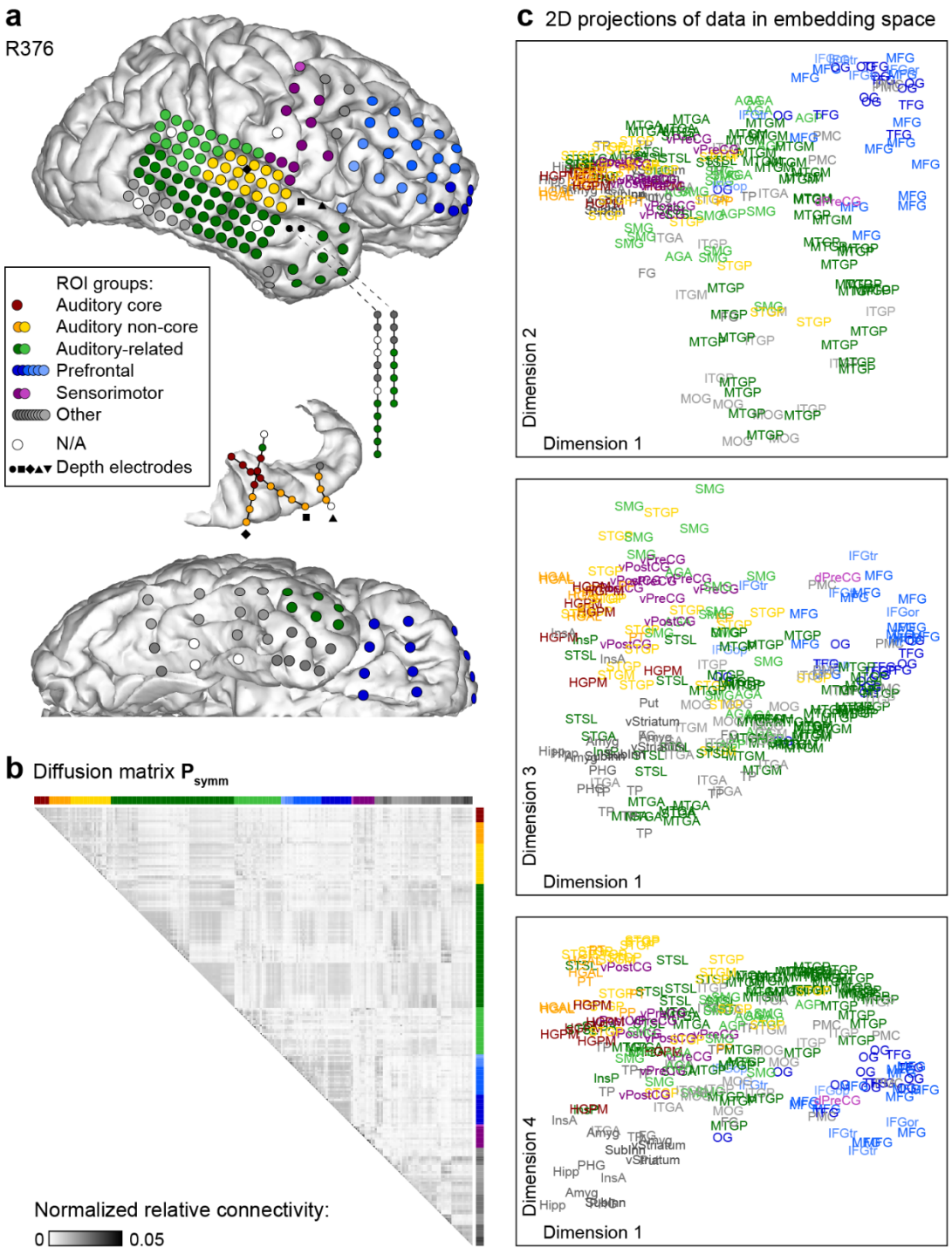


Figure 2. Functional geometry of cortical networks revealed by DME applied to gamma-band power envelope correlations in a single participant (R376). **a:** Electrode coverage. **b:** Diffusion matrix $\mathbf{P}_{symm}$. **c:** Data plotted on the same scale in the 1st and 2nd, 1st and 3rd, and 1st and 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.

Functional geometry of cortical networks

To pool data across participants with variable electrode coverage, $\mathbf{P}_{symm}$ matrices were computed at the ROI level and averaged across participants. The results for gamma-band data are shown in Figure 3a. The eigenvalue spectrum $|\lambda_i|$ of this averaged $\mathbf{P}_{symm}$ showed a clear separation between the first four and the remaining dimensions (Fig. 3a, inset), indicating that the first four dimensions of embedding space accounted for much of the community structure of the data. Indeed, these first four dimensions accounted for >80% of the diffusion distance averaged across all pairwise distances in the space, a typical measure for deciding which dimensions to retain when DME is used as a dimensionality reduction method (Coifman and Hirn, 2014).

The data are plotted in the first four dimensions of embedding space in Figure 3b (see also Supplementary Fig. 3 and Supplementary Movies 1 and 2), providing a graphical representation of the functional geometry of all sampled brain regions. Functionally related ROIs clustered together, and these clusters segregated within embedding space. For example, auditory cortical and prefrontal ROIs were at opposite ends of dimension 1, as were visual cortical (ITGP, ITGM, LinG, FG) and prefrontal ROIs. Parietal and limbic ROIs were at opposite ends of dimension 2, and auditory and visual ROIs were maximally segregated along dimension 4. By contrast, some ROIs [e.g., STGA, anterior and middle portions of middle temporal gyrus (MTGA, MTGM), middle cingulate (CingM)] were situated in the interior of the data cloud.

DME elucidates fine-scale functional organization beyond anatomical proximity

The connectivity metric employed here discards components exactly in phase between two brain regions, mitigating the influence of volume conduction (Hipp *et al.*, 2012). However, brain areas that are anatomically close to each other are often densely interconnected (Cavada et al., 2000; Jones et al., 1978; Kaas and Hackett, 1998; Kaas and Hackett, 2000; Morel et al., 1993). Thus, anatomical proximity is expected to contribute to the observed functional geometry. Overall, however, anatomical proximity explained only 14% of the variance in embedding distance (mean adjusted $r^2 = 0.14$ for regressions between anatomical and embedding Euclidean distance, calculated separately for each ROI). Anatomically adjacent ROIs that were separated in embedding space included STGA and STGM, temporal pole (TP) and the rest of the anterior temporal lobe (ATL), and InsA and InsP. Thus, the embedding representation elucidates organizational features beyond anatomical proximity.

Planum polare (PP) and posterior insula (InsP) are functionally distinct from other auditory cortical ROIs

The grouping of canonical auditory ROIs is apparent in Figure 3b, as PT, HGAL, and middle and posterior portions of the superior temporal gyrus (STGM, STGP) were all close to HGPM in

197 embedding space. One notable exception, planum polare (PP), located immediately anterior to
198 anterolateral Heschl's gyrus (HGAL), segregated from the rest of auditory cortical ROIs along
199 dimension 2 in embedding space (Fig. 3b, upper panel, lower left corner). This result is
200 consistent with PP being a higher order auditory area.

201 In contrast, InsP is a region that is anatomically distant from HGPM yet responds robustly to
202 acoustic stimuli (Zhang *et al.*, 2019), suggesting that a portion of this area could be considered
203 an auditory region (Remedios et al., 2009). For example, InsP can track relatively fast (>100 Hz)
204 temporal modulations, similar to HGPM (Steinschneider et al., 2013; Zhang *et al.*, 2019),
205 possibly due to direct inputs from the auditory thalamus. However, InsP was functionally
206 segregated from HGPM and was situated between auditory and limbic ROIs, consistent with the
207 broader role of InsP in polysensory exteroceptive processing and interoception (Craig, 2003;
208 Kuehn et al., 2016).

209

210 *Hierarchical distinction of STSU and STSL*

211 Unlike InsP and PP, STSU clustered with early auditory regions, and was significantly closer to
212 auditory cortex (core and non-core ROIs; see Fig. 1) in embedding space compared to STSL (test
213 by permutation of STSU/STSL labels, $p<0.001$). This distinction between STSL and STSU is
214 consistent with differences in their response properties reported recently (Nourski *et al.*, 2021).
215 Particularly, responses in STSL, but not STSU, were predictive of performance in a semantic
216 categorization task. Those results suggest that STSL would likely be closer in embedding space
217 to regions involved in semantic processing compared to STSU. Indeed, STSL was significantly
218 closer to ROIs reported to contribute to semantic processing [inferior frontal gyrus (IFG) pars
219 operculum/triangularis/orbitalis (IFGop, IFGtri, IFGor), TP, STGA, MTGA, MTGP, anterior and
220 posterior portions of inferior temporal gyrus (ITGA, ITGP), anterior and posterior angular gyrus
221 (AGA, AGP), supramarginal gyrus (SMG)] (Binder et al., 2009; Humphreys et al., 2015; Jackson et
222 al., 2016) compared to STSU (test by permutation of STSU/STSL labels, $p<0.001$).

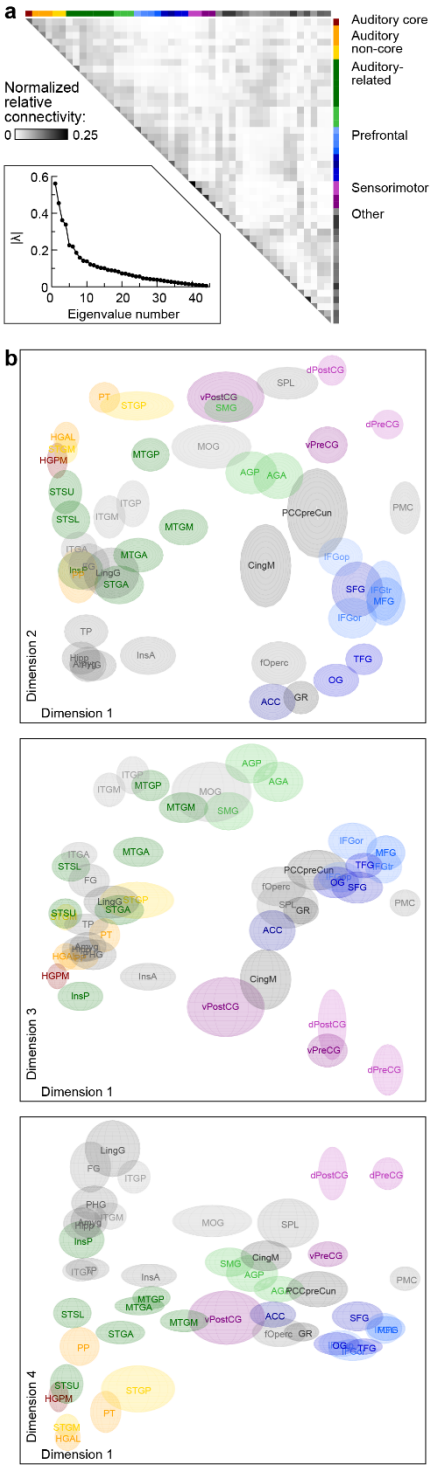


Figure 3. Summary of functional geometry of cortical networks via DME applied to gamma-band power envelope correlations. **a**: Average diffusion matrix. **Inset:** Eigenvalue spectrum. **b:** Data plotted on the same scale in the 1st and 2nd, 1st and 3rd, and 1st and 4th dimensions of embedding space (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.

Organization of ROIs outside auditory cortex

Figure 3b also characterizes the temporal and parietal ROIs outside auditory cortex that are nonetheless part of the extended auditory network, including components of the dorsal and ventral processing streams. These 'auditory-related' ROIs (shades of green in Fig. 3b), were distributed along a considerable extent of all four dimensions, consistent with functional heterogeneity of these regions and their involvement in multimodal integration (Bernstein and Liebenthal, 2014).

This heterogeneity, as well as the embedding locations of PP and STSU, suggests that the brain parcellation scheme from Figure 1 is suboptimal. Indeed, there were no quantitative criteria in this scheme for designating ROIs as 'Auditory-related' versus 'Auditory non-core'. Similarly, the 'Other' group contains a large and diverse collection of ROIs whose relationship to auditory structures and speech and language processing is unclear. To facilitate arranging these and other ROIs into functional groups or streams and develop a data-driven parcellation scheme, we turned to a quantitative hierarchical clustering approach.

Hierarchical clustering identifies mesoscale-level organizational features: ROI groups and processing streams

Hierarchical clustering applied to the first four dimensions of the embedded data shown in Figure 3 elucidated the mesoscale organization of cortical ROIs (Fig. 4) in agreement with the qualitative observations discussed above. Auditory cortical ROIs (excluding PP) formed an 'Auditory' cluster with STSU. Another major cluster (labeled 'Limbic') included ROIs traditionally considered part of the limbic system [parahippocampal gyrus (PHG), amygdala and hippocampus], as well as TP and the insula. ROIs typically considered part of the ventral and dorsal auditory streams segregated into two clusters. Additional clusters included ROIs in the ventral visual stream and those involving sensorimotor functions (labeled 'Visual' and 'Action', respectively in Fig. 4), and several clusters of prefrontal and medial cortical ROIs involved in executive function. Thus, the hierarchical clustering analysis revealed a segregation of ROIs in embedding space that aligned with known functional differentiation of brain regions. Further, we can use this analysis to expand our understanding of hierarchical relationships among clusters. For example, the 'Auditory' cluster is distinct from other clusters primarily in the temporal lobe, but is closer to the 'Limbic' cluster than 'Ventral' or 'Visual'.

In addition to these resting state recordings, most participants engaged in additional experiments investigating representation of acoustic stimuli in the brain (Nourski et al., 2017; Nourski *et al.*, 2021; Nourski et al., 2022; Steinschneider et al., 2014). We used these data to evaluate auditory responsiveness of each recording site (Fig. 4b) and compare these response profiles to the clustering results of Figure 4a. As expected, ROIs in the auditory cluster exhibited consistently high responsiveness to auditory stimuli, while visual ROIs did not. By contrast,

some clusters exhibited mixed responsiveness (e.g. InsP in the limbic cluster), possibly indicating ROIs that serve as nodes bridging auditory and other brain networks.

A new brain parcellation scheme based on these clustering results is illustrated in Figure 4c.

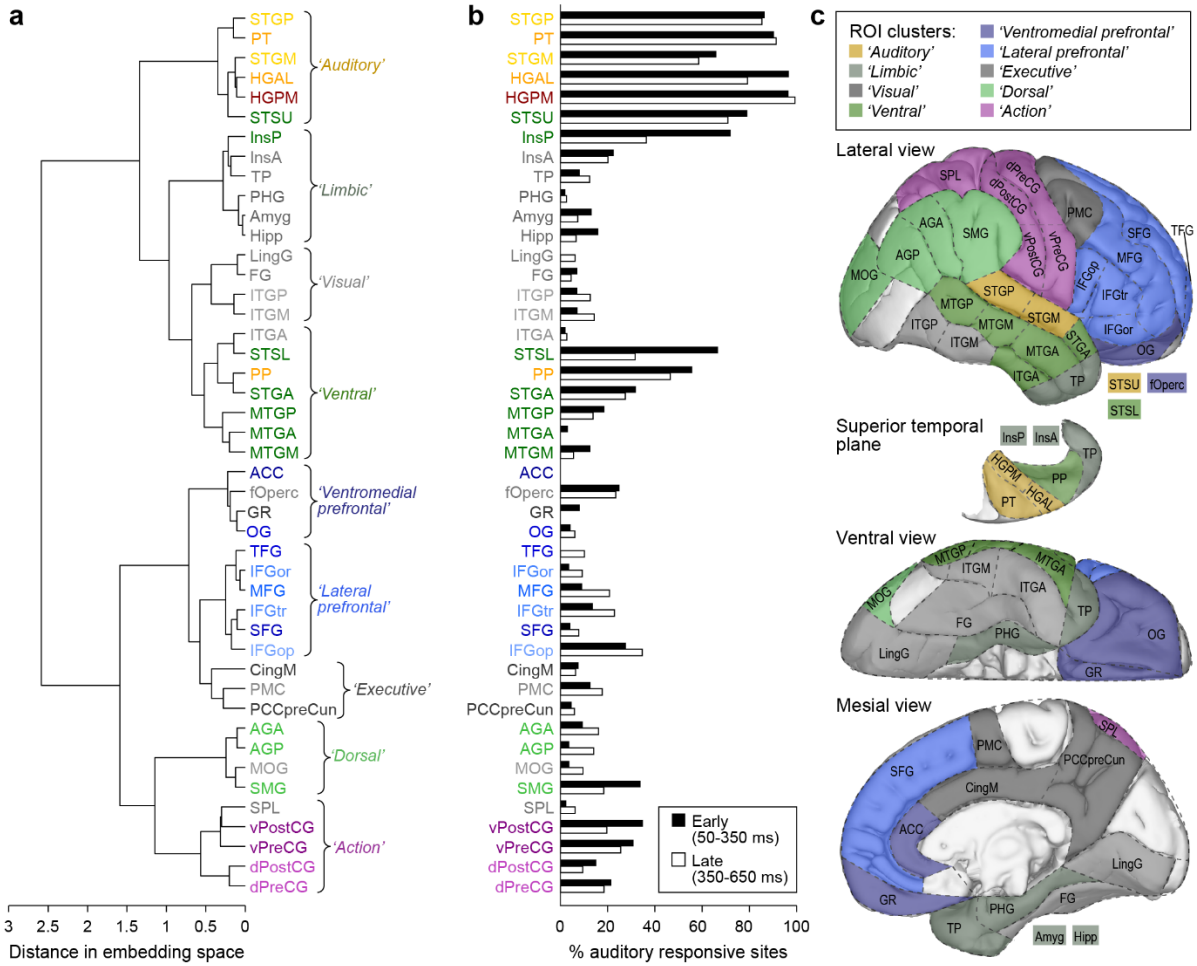


Figure 4. Hierarchical clustering of embedding data shown in Figure 3. **a:** Linkages between ROI groups identified using agglomerative clustering. b: Percentages of sites with early (50-350 ms after stimulus onset; black bars) and late (350-650 ms; white bars) high gamma responses to 300 ms monosyllabic words. **c:** Brain parcellation based on hierarchical clustering.

Embedding and hierarchical clustering in the theta band

DME applied to theta-band power envelope correlations yielded results broadly similar to the gamma band, especially in the first two dimensions of embedding space (Supplementary Figure 4). Auditory cortical ROIs other than PP clustered together and with STSU, and were maximally segregated from PFC ROIs along dimension 1. In addition, auditory-related ROIs were dispersed in embedding space, consistent with their functional heterogeneity.

Hierarchical clustering of data from theta-band power envelope correlations (Supplementary Figure 5) yielded several clusters that overlapped with those from gamma-band data. These included a cluster of auditory cortical ROIs, the bulk of lateral prefrontal cortex, and most of the 'action' cluster. Other clusters were less consistent, suggesting the temporal scale of neuronal signaling contributes to establishing distinct functional networks (Hacker et al., 2017; Keitel and Gross, 2016; Kiebel et al., 2008).

DME identifies mesoscale topological features of cortical networks

Identification of 'global hubs' within brain networks is critical for understanding their topology (Bullmore and Sporns, 2009). These nodes integrate and regulate information flow in the network by virtue of their centrality and strong connectivity, yet a precise method for identifying these hubs is yet to be established.

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. A simulated example is illustrated in Figure 5a, which depicts a 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).

We computed distance from the center of embedding space for all of the ROIs in Figure 3b. We also computed mean functional connectivity for each ROI and show in Figure 5b an overall inverse relationship between these two measures. ROIs close to the center of embedding space also exhibited strong mean connectivity, suggesting their roles as global hubs. These ROIs included MTGA, STGA, and MTGM, which all lie in the upper left quadrant of the plot >2 standard deviations from the center of the data cloud (outer dashed ellipse). ITGA, CingM, posterior cingulate/precuneus (PCC/preCun), PP, fOperc, and STSL also exhibited hub-like properties, i.e., were located in the upper left quadrant of Figure 5b. ROIs far from the center of embedding space, mostly unimodal sensory and motor regions, exhibited weak overall connectivity, consistent with their roles as spokes in the network.

In contrast to the gamma-band data, the same analysis applied to theta-band data identified CingM and ACC as two prominent global hubs, along with MTGM (Supplementary Figure 6).

These results are consistent with network organization depending on temporal scale, and suggests that mesial cortical structures regulate information flow on slower time scales. Thus, DME can identify topological features critical to information flow within cortical networks.

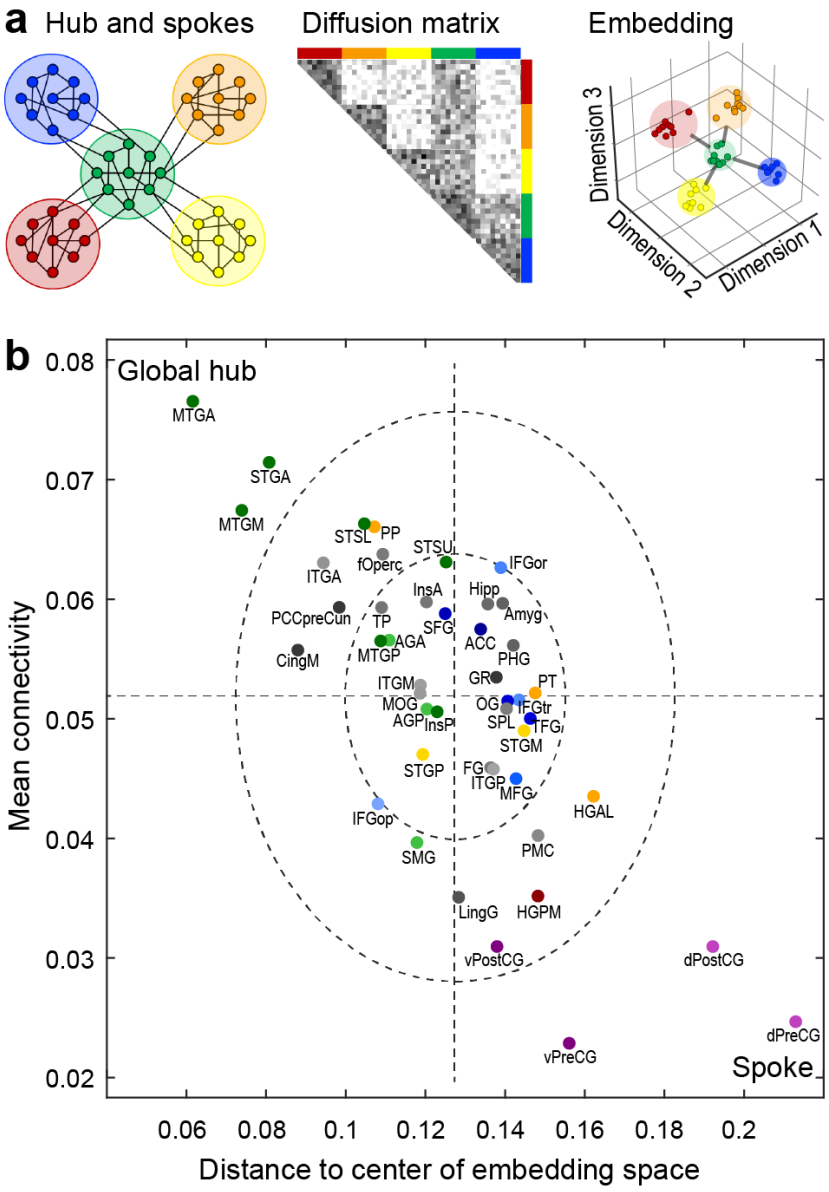


Figure 5. Identification of network hubs. **a:** Schematic example illustrating the central positioning of global hubs in embedding space. **b:** ROIs from average embedding are plotted according to their mean connectivity to the rest of the network versus their distance to the centroid of the data cloud 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.

Differences between language-dominant and non-dominant hemispheres are not specific to speech and language ROIs

On a macroscopic scale, speech and language networks are lateralized in the human brain, with nearly all right-handed and most left-handed individuals left hemisphere language-dominant (Knecht et al., 2000). However, both hemispheres are activated during speech processing (Hickok and Poeppel, 2007; Price, 2012; Schirmer et al., 2012; Turkeltaub and Coslett, 2010), and the extent to which lateralization is reflected in asymmetries in the organization of resting state auditory networks is unclear. We investigated this issue by comparing the functional geometry of cortical networks derived from participants with electrode coverage in the language-dominant (N = 24) versus non-dominant (N = 22) hemisphere. ROIs in the two hemispheres exhibited a similar functional organization in embedding space (Supplementary Fig. 7). Permutation analysis indicated that the positions of ROIs in embedding space were not significantly different between dominant and non-dominant hemispheres (all p-values > 0.05). Furthermore, there was no significant correlation between the change in position in embedding space and either early or late auditory responsiveness (early: p = 0.94; late: p = 0.86; Fig. 6a).

We also analyzed inter-ROI distances to determine whether functional interactions between ROIs were different in the two hemispheres. Pairwise inter-ROI distances in embedding space, calculated separately for dominant versus non-dominant hemisphere, were highly correlated (r = 0.88), with no obvious outliers (Fig. 6b, left panel). The data shown in Figure 6a have a slope <1, indicating that inter-ROI distances are consistently longer in the dominant hemisphere compared to the non-dominant hemisphere (p = 0.0052). This multiplicative scaling of the distances is consistent with the data occupying a larger volume in embedding space for the dominant versus non-dominant hemisphere, suggesting a greater functional heterogeneity for the language-dominant side of the brain. After accounting for this multiplicative scaling effect, following FDR correction, there were no specific inter-ROI distances that were significantly different between the two hemispheres.

When considering ROIs specifically involved in speech and language comprehension and production [PT, PP, STSL, STGP, STGM, STGA, SMG, AGA, premotor cortex (PMC), precentral gyrus (PreCG), IFGop, IFGtr] (Ardila et al., 2016; Chang *et al.*, 2015; Hickok and Poeppel, 2015), the correlation in pairwise inter-ROI distances in embedding space was also high (r = 0.90; Figure 6b). Furthermore, the data in Figure 6b exhibit a similar multiplicative scaling as observed for all the ROIs shown in Figure 6a. Indeed, the slope for the data in Figure 6b was indistinguishable from the slope for the data in Figure 6a (p = 0.92). Thus, hemispheric asymmetry of functional organization specific to speech and language networks was not detectable in RS connectivity.

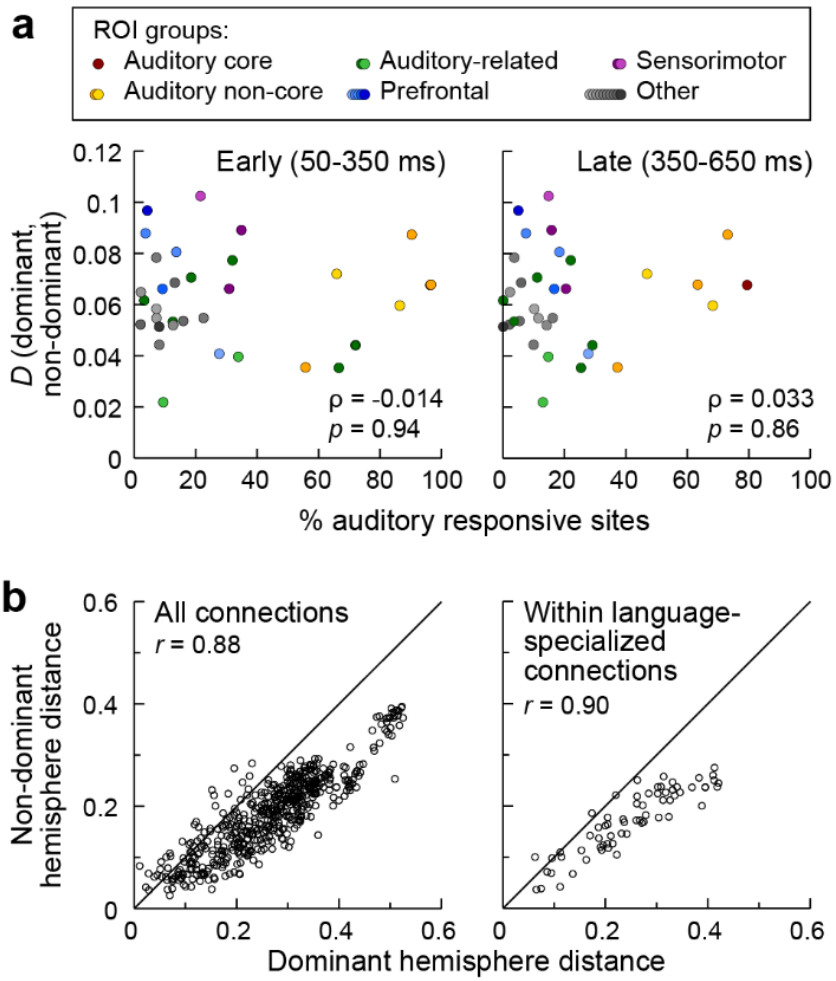


Figure 6. RS connectivity is symmetric between hemispheres. Inter-ROI distances in embedding space for non-dominant versus dominant hemisphere participants. **a:** Comparison between the change in position in embedding space from dominant to non-dominant hemisphere and the auditory responsiveness of individual ROIs. Two-tailed Spearman's rank tests did not reveal a significant correlation between ROI asymmetry and percentage of either early or late auditory responsive sites within the ROI (left and right panel, respectively). **b:** 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 bands

Other bands (alpha, beta, high gamma) produced similar embeddings to those from gamma and theta. Inter-ROI distances were highly similar for adjacent bands ($r \geq 0.82$), and even for non-adjacent bands ($r \geq 0.67$; Supplementary Fig. 8). Thus, DME identified overall, rather than band-specific, organizational features of cortical networks.

However, a particular band might be preferred if it produced narrower estimation margins in the functional geometry. An overall relative uncertainty was calculated as the correlation between inter-ROI embedding distances in the original data versus bootstrapped data. Correlation values were uniformly high across bands (r = 0.91, 0.85, 0.87, 0.88, and 0.86 for high gamma, gamma, beta, alpha, and theta, respectively). These analyses suggest that DME offers a robust approach to exploring functional geometry.

Comparison to embeddings derived from RS-fMRI data

So far, we've presented results at multiple spatial scales based on intracranial electrophysiology. However, these intracranial recordings sample the brain non-uniformly and sparsely as dictated by clinical considerations. This feature presents problems at two spatial scales: first, cortical regions are not sampled uniformly (with some not sampled at all). Second, ROIs are not sampled uniformly across their volume. To examine the impact of these sampling issues, we compared iEEG-based DME to DME applied to RS-fMRI data available in a subset of ten participants.

We first tested 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 correlated (Fig. 7c), with highest correlations for gamma- and high gamma-band envelopes ($r >$ 0.45; Fig. 7d, line and symbols).

The analysis presented in Figure 7 provide a context for using fMRI data to address questions regarding the effects of limited, non-uniform sampling. We used a standard parcellation scheme developed for fMRI data (Schaefer-Yeo 400 ROIs; (Schaefer et al., 2018)) rather than the iEEG parcellation scheme introduced in Figure 1.

The first question we addressed was the effect of non-uniformly sampling only a subset of brain regions. For each participant, embeddings were derived from RS-fMRI connectivity matrices computed from all cortical ROIs (Fig. 8a, "Full fMRI", first column). From these embeddings, we selected only points in embedding space corresponding to ROIs sampled with iEEG (Fig. 8a, "Full fMRI (iEEG subset)", second column). We also computed embeddings for each subject from

405 only the fMRI ROIs sampled with iEEG in that subject ["Partial fMRI (ROI level)", Fig. 8a, 3rd
406 column]. We compared these embeddings to the "Full fMRI (iEEG subset)" embeddings by
407 computing the correlation between inter-ROI distances (Fig. 8b). Although the scale of the
408 embeddings was different for the full fMRI versus partial fMRI data (because the number of
409 dimensions was different), the two were highly correlated (median r = 0.90; Fig. 8c). Thus,
410 embeddings constructed from the portion of the brain sampled by iEEG were quite similar to
411 embeddings derived from the whole brain.

412 The second question we addressed was the effect of representing an entire ROI by sparse
413 sampling with a limited number of electrodes. We computed embeddings from the voxel
414 averages across entire ROIs in each participant ["Partial fMRI (ROI level)", Fig. 8a, 3rd column]
415 and from averages of the voxels in grey-matter spheres around iEEG recording sites ["Partial
416 fMRI (site level)", Fig. 8a, rightmost column]. ROI- and site-level embedding distances were
417 strongly correlated (median r = 0.65; Fig. 8c).

418 Thus, sparse sampling within an ROI had a greater impact on estimates of functional geometry
419 than limited sampling of the complete set of ROIs. Overall, however, ROIs were faithfully
420 represented in embedding space even when DME was based on a small number of locations
421 within ROIs. Taken together, these results indicate broad consistency between functional
422 organization derived from iEEG and fMRI and the robustness of this approach to sparse
423 sampling afforded by iEEG recordings.

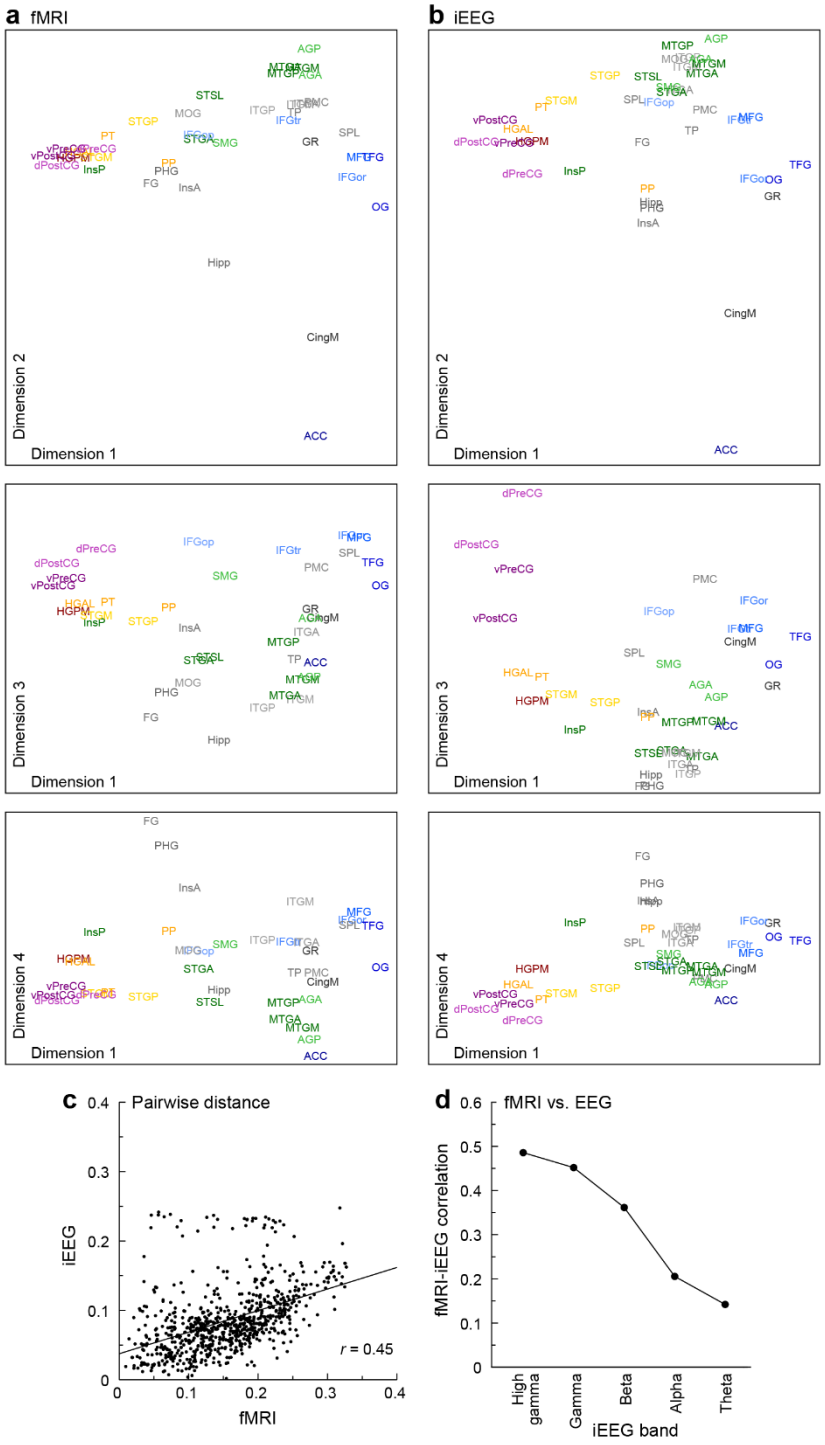


Figure 7. Comparison of iEEG and fMRI connectivity data in embedding space. **a:** Participant-averaged embeddings for iEEG (gamma band power envelope correlations). **b**: Participant-averaged embeddings for fMRI. Scale bar: 0.1. **c:** Inter-ROI embedding distances computed from the data in **a** and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

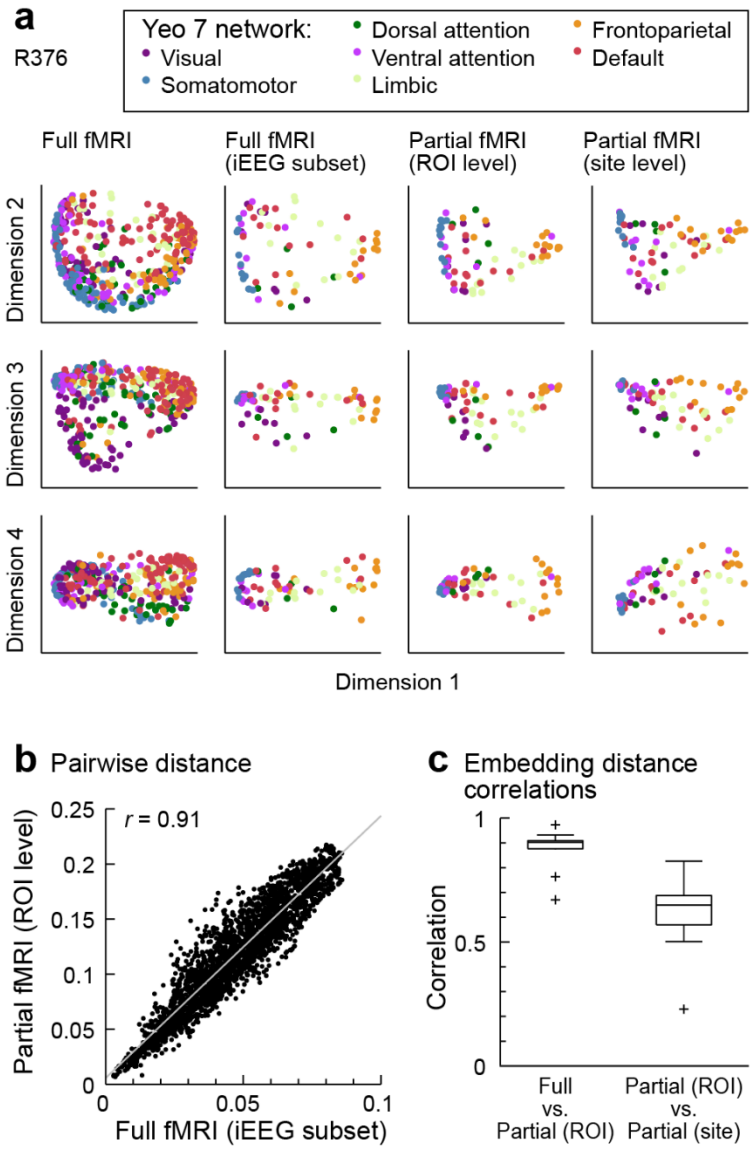


Figure 8. Comparison of embeddings derived from full fMRI connectivity matrices and connectivity matrices computed using only ROIs sampled with iEEG. **a:** Data in the first four dimensions of embedding space for a single participant. Shown are embeddings of all derived from the full RS-fMRI connectivity matrix (1st column); the subset of the data points in the 1st column corresponding to ROIs sampled via iEEG (2nd column); and embeddings derived from connectivity matrices including *only* the ROIs sampled via iEEG, calculated by averaging across the entire ROI (3rd column), and calculated based on the specific recording sites in that participant (4th column). **b:** Comparison of embedding distances calculated from the full fMRI embedding (i.e, data in **a**, 2nd column) versus distances calculated from the partial fMRI embedding (i.e., data in **a**, 3rd column). **c:** Summary across participants of distance correlations between full fMRI embeddings versus partial embeddings calculated based on the entire ROI (*left*, "Full vs. Partial (ROI)") and between partial embeddings calculated based on the entire ROI versus those calculated based on recording sites [*right*, "Partial (ROI) vs. Partial (site)"].

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 at multiple spatial scales. We demonstrate methodology for testing specific hypotheses at each of these scales using DME. We also generate data-driven hypotheses for study using future data sets.

Fine scale: Organization of auditory cortex

At a fine spatial scale, previous work in macaque has defined over a dozen auditory cortical fields based on cytoarchitectonics, connectivity, and response properties (Hackett et al., 2001). By contrast, there is no consensus on how auditory cortex is organized in humans, with multiple candidate parcellations based on cytoarchitectonics, tonotopy or myeloarchitecture (Barton et al., 2012; Hackett, 2015; Moerel et al., 2014; Woods et al., 2010). Our results contribute to this body of knowledge by showing that several superior temporal ROIs including core auditory cortex (HGPM) and putative auditory belt and parabelt areas (PT, HGAL, STGP, STGM) (Hackett, 2015; Moerel *et al.*, 2014) cluster together in embedding space. Thus, in spite of their diversity in processing of specific features of acoustic signals, these ROIs are positioned at a similar level in the auditory processing hierarchy. Proximity of STGP and STGM to HGPM in embedding space is consistent with previous studies that interpret these regions as relatively early non-core auditory cortex (Hamilton et al., 2021; Howard et al., 2000; Nourski et al., 2014). By contrast, PP is anatomically close and connected to HGPM (Upadhyay et al., 2008), yet it is distinguished among auditory cortical regions for its syntactic-level language processing (Friederici *et al.*, 2000) and its preferential activation by music, which has a strong affective component (Angulo-Perkins *et al.*, 2014). This functional differentiation is reflected in its segregation from the auditory cluster in embedding space.

Fine scale: 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 (Abrams *et al.*, 2020; Beauchamp, 2015; Chang *et al.*, 2015; Hickok, 2009; Price, 2012; Venezia *et al.*, 2017). Previous studies have shown that STSU and STSL differ in cytoarchitecture (Zachlod et al., 2020) and have distinct responses to speech (Belin et al., 2000; Deen et al., 2015; Leaver and Rauschecker, 2010; Wilson et al., 2018). A recent iEEG study demonstrated enhanced, shorter-latency, responses to speech syllables in STSU compared to STSL (Nourski *et al.*, 2021). STSU is traditionally not considered part of canonical auditory cortex (but see (Woods *et al.*, 2010)), 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 (Nourski *et al.*, 2021). The regions specifically involved in semantic processing is a current topic of debate, with multiple competing models (Binder *et al.*, 2009; Humphreys *et al.*, 2015; Jackson *et al.*, 2016; Lambon Ralph *et al.*, 2017). We defined a list of semantic processing regions by combining across these models.-STSL was positioned closer in embedding space to these regions compared to STSU. Taken together, the results firmly place STSU and STSL at different levels of the auditory cortical hierarchy.

Mesoscale: Functional and theoretical framework of a limbic auditory pathway

Multiple lines of evidence support a pathway linking auditory cortical and limbic structures (Kahn et al., 2008; Michelmann et al., 2021; Rocchi et al., 2021; Wang et al., 2016) that subserves auditory memory (Kumar *et al.*, 2021; Kumar *et al.*, 2016; Munoz-Lopez *et al.*, 2010) and affective sound processing (Fruhholz et al., 2016). The data presented here contribute to our understanding of this pathway. Clustering analysis identified a set of ROIs including structures classically labeled as limbic (PHG, Amy, Hipp) as well as insula (InsP, InsA) and TP positioned close to the auditory cluster in embedding space (Fig. 4). This suggests a close functional relationship that could form the basis for a limbic stream. InsP, with strong auditory responsiveness and overlapping response properties with HGPM, is likely involved in the transformation of auditory information in auditory cortex to affective representations in InsA (Zhang *et al.*, 2019). Thus InsP could serve as critical linking node between auditory and limbic structures.

TP is involved in semantic processing (Friederici *et al.*, 2000; Lambon Ralph *et al.*, 2017) and auditory memory (Munoz-Lopez et al., 2015), in particular the representation and retrieval of memories for people, social language, and behaviors ('social knowledge') (Olson et al., 2013). Tight clustering of TP with limbic ROIs in embedding space is consistent with its previously reported functional association with limbic cortex (Chanes and Barrett, 2016; Mesulam, 2000), with which TP shares key features of laminar cytoarchitecture and strong connectivity (Maller et al., 2019). We suggest that the organization depicted in Figures 3 and 4, combined with evidence for bidirectional information sharing between auditory cortex and limbic areas, merits the identification of a third auditory processing stream alongside the dorsal and ventral streams (Hickok, 2012; Rauschecker and Scott, 2009). This 'limbic stream' would underlie auditory contributions to affective and episodic memory processing.

Mesoscale: Ventral and dorsal streams linking auditory and frontal cortex

Current models of speech and language processing posit the existence of ventral and dorsal processing streams linking non-core auditory cortex with PMC and inferior frontal gyrus via several distinct anatomical pathways encompassing temporal, parietal, and frontal cortex

(Chang *et al.*, 2015; Friederici, 2012; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Despite substantial experimental evidence supporting these models, there is a lack of consensus on the specific functions subserved by the two streams. For example, the dorsal stream has been envisioned to subserve spatial processing ("where" (Rauschecker and Scott, 2009)), sensorimotor integration ("how" (Hickok and Poeppel, 2007)), and syntactic processing (Friederici, 2012). There is a parallel debate about the specific cortical regions comprising the two streams.

As broadly predicted by these models, temporal and parietal ROIs segregated in embedding space in the analysis presented here (Fig. 3b, 4). We observed a cluster that included STSL, middle and inferior temporal gyrus ROIs, in conformity with the ventral auditory stream proposed by Hickok and Poeppel (Hickok and Poeppel, 2007) and Friederici (Friederici, 2012). By contrast, the cluster that included SMG, AGP, and AGA aligned with the dorsal processing stream as proposed by Rauschecker and Scott (Rauschecker and Scott, 2009). Association of FG and MOG with the ventral and dorsal clusters, respectively, likely 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 (Margulies *et al.*, 2016). 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 (Friederici, 2012; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009).

Mesoscale: Network hubs

Hubs in brain networks play a critical role in integrating distributed neural activity (Bullmore and Sporns, 2009; van den Heuvel and Sporns, 2013). In the present analysis, global hubs were characterized by their central location within embedding space and high mean connectivity (Fig. 5). In the gamma band, 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 (Abel et al., 2015; Lambon Ralph *et al.*, 2017; Simmons and Martin, 2009) and playing a central role in semantic processing and social memory (Lambon Ralph *et al.*, 2017; Olson *et al.*, 2013; Patterson et al., 2007). 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 (Scott et al., 2000; Spitsyna et al., 2006), but the damage is often more widespread and can include MTGM (Gorno-Tempini et al., 2004).

Cingulate cortical ROIs (CingM, ACC) were identified as hubs in theta-band data. These areas are described as transmodal and are active during a wide array of emotional and cognitive

processes (Mesulam, 1998; Rolls, 2019), both consistent with their previous characterization as network hubs (van den Heuvel and Sporns, 2013). The identification of hubs specific to each frequency band supports the model in which the temporal scale of communication in the brain supports distinct functional networks (Hacker *et al.*, 2017; Keitel and Gross, 2016; Kiebel *et al.*, 2008).

Unlike other ATL structures, TP does not appear as a global hub in in either gamma or theta bands (Fig. 5b, Supplementary Figure 6). 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 (Binney et al., 2012; Lambon Ralph *et al.*, 2017; Makris et al., 2009).

Macroscale: 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 (Binder et al., 2000; Cogan et al., 2014; de Heer et al., 2017). We found evidence for hemispheric differences in cortical functional organization based on analysis of all sampled brain regions, with inter-ROI distances being systematically greater in embedding space for the language-dominant hemisphere (Fig. 6b). This is consistent with greater inter-regional heterogeneity in that hemisphere compared to the non-dominant side. Importantly, ROIs involved in speech and language processing did show any additional asymmetry (Fig. 6b), nor was the difference in position in embedding space related to auditory responsiveness (Fig. 6a). Recent studies that identified interhemispheric differences in RS connectivity for the STS (Abrams *et al.*, 2020) and semantic networks more broadly (Gonzalez Alam et al., 2021) may reflect this broader asymmetry observed here. Our are also consistent with a recent fMRI study showing RS connectivity patterns in lateral temporal cortex that were comparable between left and right hemispheres (Jackson *et al.*, 2018). This does not exclude the possibility of asymmetries specific to auditory regions emerging during sensory tasks, for example reflecting hemispheric biases in spectral and temporal processing (Hickok and Poeppel, 2007; 2015).

Caveats & limitations

A key concern regarding all human iEEG studies is that participants may not be representative of a healthy population. In the present study, results were consistent across participants despite differences in seizure disorder histories, medications, and seizure foci, and aligned with results obtained previously in healthy participants (Margulies *et al.*, 2016). Another caveat is that our dataset, however extensive, did not sample the entire brain, and it was not possible to

591 infer connectivity with unsampled regions. To address this, we applied DME analysis to fMRI
592 data to establish that the organization of ROIs in embedding space was robust to the exclusion
593 of unsampled ROIs. Although there was a greater effect of sparse, non-uniform sampling within
594 an ROI, there was still considerable similarity in functional organization to embeddings derived
595 from averages across the entire ROI.

596 While subcortical structures (e.g., thalamus) that link sensory and higher order networks
597 (Sherman and Guillery, 2011) were not sampled, the functional organization presented here
598 was likely influenced indirectly by thalamo-cortical pathways (Hamilton *et al.*, 2021; Hu, 2003).
599 Previous fMRI studies of RS networks focused exclusively on cortical ROIs and did not consider
600 the role of the thalamus and other subcortical structures. Despite this limitation, these studies
601 have yielded valuable insights into the functional organization of the human cortical networks
602 (Biswal *et al.*, 2010; Seitzman et al., 2019).

603

604 *Concluding remarks and future directions*

605 This study extends the DME approach to characterize functional relationships between cortical
606 regions investigated using iEEG recordings. These data help resolve several outstanding issues
607 regarding the functional organization of human auditory cortical networks and stress the
608 importance of a limbic pathway complementary to the dorsal and ventral streams. These
609 results lay the foundation for future work investigating network organization during active
610 speech and language processing. While the current work focused on auditory cortical networks,
611 this approach can be readily generalized to advance our understanding of changes in brain
612 organization during sleep and anesthesia, disorders of consciousness, as well as reorganization
613 of cortical functional geometry secondary to lesions.

Online Methods

Participants

The study was carried out in 49 neurosurgical patients (22 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 aligned with best practices recently aggregated in (Feinsinger et al., 2022) and were approved by the University of Iowa Institutional Review Board and the National Institutes of Health; 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 (42 out of 49); six participants were left-handed, and one had bilateral handedness. The majority of participants (35 out of 49) 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 participant with bilateral dominance, and the remaining two participants who did not undergo Wada test and who were left-handed were not included in the analysis of hemispheric asymmetry in Figure 6.
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.

iEEG recordings. iEEG recordings were obtained using either subdural and depth electrodes, or depth electrodes alone, based on clinical indications. Electrode arrays were manufactured by Ad-Tech Medical (Racine, WI). Subdural arrays, implanted in 36 participants out of 46, consisted of platinum-iridium discs (2.3 mm diameter, 5-10 mm inter-electrode distance), embedded in a silicon membrane. Stereotactically implanted depth arrays included between 4 and 12 cylindrical contacts along the electrode shaft, with 5-10 mm inter-electrode distance. A subgaleal electrode, placed over the cranial vertex near midline, was used as a reference in all participants. All electrodes were placed solely on the basis of clinical requirements, as determined by the team of epileptologists and neurosurgeons (Nourski and Howard, 2015).

No-task RS data were recorded in the dedicated, electrically shielded suite in The University of Iowa Clinical Research Unit while the participants lay in the hospital bed. RS data were collected 6.4 +/- 3.5 days (mean ± standard deviation; range 1.5 – 20.9) after electrode implantation surgery. In the first 15 participants (L275 through L362), data were recorded using a TDT RZ2 real-time processor (Tucker-Davis Technologies, Alachua, FL). In the remaining 34 participants (R369 through L585), data acquisition was performed using a Neuralynx Atlas System (Neuralynx Inc., Bozeman, MT). Recorded data were amplified, filtered (0.1–500 Hz bandpass, 5 dB/octave rolloff for TDT-recorded data; 0.7–800 Hz bandpass, 12 dB/octave rolloff for Neuralynx-recorded data) and digitized at a sampling rate of 2034.5 Hz (TDT) or 2000 Hz (Neuralynx). The durations of recordings were 13 +/- 11 min. In all but two participants, recording durations were between 10 and 22 min.; in one participant duration was 6 min., and in one participant the duration was 81 min.

Data analysis

Anatomical reconstruction and ROI parcellation. Localization of recording sites and their assignment to ROIs relied on post-implantation T1-weighted anatomical MRI and post-implantation computed tomography (CT). All images were initially aligned with pre-operative T1 scans using linear coregistration implemented in FSL (FLIRT) (Jenkinson et al., 2002). Electrodes were identified in the post-implantation MRI as magnetic susceptibility artifacts and in the CT as metallic hyperdensities. Electrode locations were further refined within the space of the pre-operative MRI using three-dimensional non-linear thin-plate spline warping (Rohr et al., 2001), which corrected for post-operative brain shift and distortion. The warping was constrained with 50-100 control points, manually selected throughout the brain, which were visually aligned to landmarks in the pre- and post-implantation MRI.

To pool data across participants, the dimensionality of connectivity matrices was reduced by assigning electrodes to one of 58 ROIs organized into 6 ROI groups (see Fig. 1; Supplementary Table 2, 3) based upon anatomical reconstructions of electrode locations in each participant.

For subdural arrays, ROI assignment was informed by automated parcellation of cortical gyri (Destrieux et al., 2010; Destrieux et al., 2017) as implemented in the FreeSurfer software package. For depth arrays, it was informed by MRI sections along sagittal, coronal, and axial planes. For recording sites in Heschl's gyrus, delineation of the border between core auditory cortex adjacent non-core areas (HGPM and HGAL, respectively) was performed in each participant using physiological criteria (Brugge et al., 2009; Nourski et al., 2016). Specifically, recording sites were assigned to HGPM if they exhibited phase-locked (frequency-following) responses to 100 Hz click trains and if the averaged evoked potentials to these stimuli featured short-latency (<20 ms) peaks. Such response features are characteristic for HGPM and are not present within HGAL (Brugge *et al.*, 2009). Additionally, correlation coefficients between average evoked potential waveforms recorded from adjacent sites were examined to identify discontinuities in response profiles along Heschl's gyrus that could be interpreted as reflecting a transition from HGPM to HGAL. Superior temporal gyrus was subdivided into posterior and middle non-core auditory cortex ROIs (STGP and STGM), and auditory-related anterior ROI (STGA) using the transverse temporal sulcus and ascending ramus of the Sylvian fissure as macroanatomical boundaries. The insula was subdivided into posterior and anterior ROIs, with the former considered within the auditory-related ROI group (Zhang *et al.*, 2019). Middle and inferior temporal gyrus were each divided into posterior, middle, and anterior ROIs by diving the gyrus into three approximately equal-length thirds. Angular gyrus was divided into posterior and anterior ROIs using the angular sulcus as a macroanatomical boundary. Anterior cingulate cortex was identified by automatic parcellation in FreeSurfer and was considered as part of the prefrontal ROI group, separately from the rest of the cingulate gyrus. Postcentral and precentral gyri were each divided into ventral and dorsal portions using the y_{MNI} coordinate (see below) of 40 mm as a boundary. Recording sites identified as seizure foci or characterized by excessive noise, and depth electrode contacts localized to the white matter or outside brain, were excluded from analyses and are not listed in Supplementary Table 2. Electrode coverage was largely restricted to a single hemisphere in individual participants, and contacts on the contralateral hemisphere were excluded from analysis (and are not listed in Supplementary Table 2) such that all connections represent intra-hemisphere functional connectivity.

Preprocessing of fMRI data. Standard preprocessing was applied to the RS-fMRI data acquired in the pre-implantation scan using FSL's FEAT pipeline, including spatial alignment and nuisance regression. White matter, cerebrospinal fluid and global ROIs were created using deep white matter, lateral ventricles and a whole brain mask, respectively. Regression was performed using the time series of these three nuisance ROIs as well as 6 motion parameters (3 rotations and 3 translations) and their derivatives, detrended with second order polynomials. Temporal bandpass filtering was 0.008–0.08 Hz. Spatial smoothing was applied with a Gaussian kernel (6 mm full-width at half maximum). The first two images from each run were discarded. Frame censoring was applied when the Euclidean norm of derivatives of motion parameters exceeded 0.5 mm (Power et al., 2012). All runs were processed in native EPI space, then the residual data were transformed to MNI152 and concatenated.

Preprocessing of iEEG data. Analysis of iEEG data was performed using custom software written in MATLAB Version 2020a programming environment (MathWorks, Natick, MA, USA). After initial rejection of recording sites identified as seizure foci, several automated steps were taken to exclude recording channels and time intervals contaminated by noise. First, channels were excluded if average power in any frequency band [broadband, delta (1-4 Hz), theta (4-8 Hz), alpha (8-13Hz), beta (13-30 Hz), gamma (30-50 Hz), or high gamma (70-110 Hz); 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 (Schaefer *et al.*, 2018) 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, envelope correlations were estimated within 60-second data segments using orthogonalized band power envelope correlations as in (Hipp *et al.*, 2012), except time-frequency decomposition was performed using the demodulated band transform (Kovach and Gander, 2016) rather than wavelets. This measure avoids artifacts due to volume conduction by discounting connectivity near zero phase lag. 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. Lastly, correlations were averaged across segments.

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 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) 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; (Coifman and Hirn, 2014; Coifman *et al.*, 2005)).

DME consists of mapping the recording sites into an embedding space using an eigendecomposition of **P**,

$$\Psi^{(t)}(x_i) = [\lambda_1{}^t\psi_1(x_i), \lambda_2{}^t\psi_2(x_i), \ldots, \lambda_M{}^t\psi_M(x_i)]^{\mathrm{T}} ,$$

where ψ_j are the eigenvectors of **P**. The parameter t is the number of time steps in that random walk; larger values of t shift focus from local to global features of the data. Here, we present data for $t = 1$. In the analyses presented here, **K** is a matrix of orthogonalized power envelope correlations transformed by applying cosine similarity (Margulies *et al.*, 2016).

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 (Coifman and Hirn, 2014). 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 (Coifman and Hirn, 2014).

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 $\mathbf{P_{symm}}$ (Coifman and Hirn, 2014). 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 **P** or $\mathbf{P}_{\mathbf{symm}}$. The inflection point k_{infl} was identified algorithmically (Satopaa et al., 2011), and the number of dimensions retained set equal to $k_{infl} - 1$.

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. To compare anatomical distance and distance in embedding space, inter-ROI anatomical distances were calculated for each participant by computing the centroid of each ROI in MNI space, then calculating Euclidean distances between centroids, followed by averaging distances across participants.

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 that 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 (Binder *et al.*, 2009; Humphreys *et al.*, 2015; Jackson *et al.*, 2016; Lambon Ralph *et al.*, 2017)), 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 groupings) would not have had sufficient statistical power under appropriate corrections for multiple comparisons. Though different choices could have been made for inclusion in the "semantic processing" category, exchanging one or two of these ROIs would not strongly influence the average distance in a group of twelve ROIs.

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

Auditory responsiveness. In a subset of 37 participants, auditory responsiveness was evaluated as percentage of sites within each ROI that exhibited high gamma responses to monosyllabic word stimuli. The stimuli were 300 ms words "cat", "dog", "five", "ten", "red", "white", presented in semantic categorization and tone target detection tasks (Nourski *et al.*, 2017; Nourski *et al.*, 2021; Nourski *et al.*, 2022; Steinschneider *et al.*, 2014). Mean high gamma (70-110 Hz) power within early (50 to 350 ms) and late (350 to 650 ms) poststimulus time windows was compared with that in a prestimulus window (-200 to -100 ms). Significance of high gamma responses was established at a $p = 0.05$ level using one-tailed Mann-Whitney *U* tests with false discovery rate correction.

Comparing language dominant/non-dominant hemispheres. To test for differences in functional geometry between language dominant and non-dominant hemispheres, two measures were considered: differences in the location of individual ROIs in embedding space, and different pairwise distances between ROIs in embedding space. To calculate differences in location of individual ROIs, dominant/non-dominant average embeddings were mapped to a common space (from an embedding using the average across all participants regardless of language dominance) using the change of basis operator. The language-dominant location difference for a specific ROI was calculated as the Euclidean distance between the two locations of each ROI in this common space. To examine whether there was a consistent relationship between hemispheric asymmetry in a given ROI's location in embedding space and the percentage of either early or late auditory responsive sites within that ROI, two-tailed Spearman's rank tests were used. To calculate differences in pairwise distances between ROIs, Euclidean distances were calculated in embedding space for each hemisphere and then subtracted to obtain a difference matrix. To determine whether the differences in location or pairwise distances were larger than expected by chance, random permutations of the dominant/non-dominant labels were used to generate empirical null distributions. Since this approach produces a *p*-value for every pair of connections, *p*-values were adjusted using false discovery rate (FDR) to account for multiple comparisons.

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

Author contributions

Conceptualization: M.I.B., K.V.N.

Methodology: M.I.B., B.M.K., A.D.B., J.E.B., C.K.K., M.S., K.V.N.

Software: M.I.B., B.M.K, D.B.G., D.I.C., C.K.K.

Validation: B.M.K.

Formal Analysis: M.I.B., B.M.K, D.B.G., D.I.C., K.V.N.

Investigation: H.K., K.V.N.

Data Curation: B.M.K., C.K.K., J.E.B., H.K., K.V.N.

Writing – Original Draft: M.I.B., B.M.K., K.V.N.

Writing – Review & Editing: M.I.B., B.M.K., D.B.G., D.I.C., A.D.B., J.E.B., C.K.K., M.S., K.V.N.

Visualization: B.M.K., K.V.N.

Supervision: M.I.B., K.V.N.

Project Administration: K.V.N.

Funding Acquisition: M.I.B., K.V.N.

Declaration of interest

The authors declare no competing interests.

Data and code availability

Software and data used to generate figures are freely available at https://zenodo.org/record/7200024 or DOI 10.5281/zenodo.7200024 . Complete data set is available via a request to the Authors pending establishment of a formal data sharing agreement and submission of a formal project outline. 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 $k(x_i,x_j)$ as the cosine similarity between functional connectivity of nodes x_i and 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 $p(x_i,x_j)=0$, **Pv** = **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 i^{th} 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 (Coifman and Hirn, 2014):

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

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 $<\mathbf{a},\mathbf{b}>$ is the inner product of **a** and **b**.

1022 REFERENCES CITED
1023
1024 Abel, T.J., Rhone, A.E., Nourski, K.V., Kawasaki, H., Oya, H., Griffiths, T.D., Howard, M.A., 3rd, and Tranel,
1025 D. (2015). Direct physiologic evidence of a heteromodal convergence region for proper naming in human
1026 left anterior temporal lobe. J Neurosci *35*, 1513-1520. 10.1523/JNEUROSCI.3387-14.2015.

1027 Abrams, D.A., Kochalka, J., Bhide, S., Ryali, S., and Menon, V. (2020). Intrinsic functional architecture of
1028 the human speech processing network. Cortex *129*, 41-56. 10.1016/j.cortex.2020.03.013.

1029 Angulo-Perkins, A., Aube, W., Peretz, I., Barrios, F.A., Armony, J.L., and Concha, L. (2014). Music listening
1030 engages specific cortical regions within the temporal lobes: differences between musicians and non-
1031 musicians. Cortex *59*, 126-137. 10.1016/j.cortex.2014.07.013.

1032 Ardila, A., Bernal, B., and Rosselli, M. (2016). How Localized are Language Brain Areas? A Review of
1033 Brodmann Areas Involvement in Oral Language. Arch Clin Neuropsychol *31*, 112-122.
1034 10.1093/arclin/acv081.

1035 Barton, B., Venezia, J.H., Saberi, K., Hickok, G., and Brewer, A.A. (2012). Orthogonal acoustic dimensions
1036 define auditory field maps in human cortex. Proc Natl Acad Sci U S A *109*, 20738-20743.
1037 10.1073/pnas.1213381109.

1038 Beauchamp, M.S. (2015). The social mysteries of the superior temporal sulcus. Trends Cogn Sci *19*, 489-
1039 490. 10.1016/j.tics.2015.07.002.

1040 Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory
1041 cortex. Nature *403*, 309-312.

1042 Bernstein, L.E., and Liebenthal, E. (2014). Neural pathways for visual speech perception. Front Neurosci
1043 *8*, 386. 10.3389/fnins.2014.00386.

1044 Binder, J.R., Desai, R.H., Graves, W.W., and Conant, L.L. (2009). Where is the semantic system? A critical
1045 review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex *19*, 2767-2796.
1046 10.1093/cercor/bhp055.

1047 Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., and Possing, E.T.
1048 (2000). Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex *10*, 512-528.
1049 10.1093/cercor/10.5.512.

1050 Binney, R.J., Parker, G.J., and Lambon Ralph, M.A. (2012). Convergent connectivity and graded
1051 specialization in the rostral human temporal lobe as revealed by diffusion-weighted imaging
1052 probabilistic tractography. J Cogn Neurosci *24*, 1998-2014. 10.1162/jocn_a_00263.

1053 Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S.,
1054 Buckner, R.L., Colcombe, S., et al. (2010). Toward discovery science of human brain function. Proc Natl
1055 Acad Sci U S A *107*, 4734-4739. 10.1073/pnas.0911855107.

1056 Brugge, J.F., Nourski, K.V., Oya, H., Reale, R.A., Kawasaki, H., Steinschneider, M., and Howard, M.A., 3rd
1057 (2009). Coding of repetitive transients by auditory cortex on Heschl's gyrus. J Neurophysiol *102*, 2358-
1058 2374. 10.1152/jn.91346.2008.

1059 Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and
1060 functional systems. Nat.Rev.Neurosci. *10*, 186-198. nrn2575 [pii];10.1038/nrn2575 [doi].

1061 Cavada, C., Company, T., Tejedor, J., Cruz-Rizzolo, R.J., and Reinoso-Suarez, F. (2000). The anatomical
1062 connections of the macaque monkey orbitofrontal cortex. A review. Cereb Cortex *10*, 220-242.
1063 10.1093/cercor/10.3.220.

1064 Chanes, L., and Barrett, L.F. (2016). Redefining the Role of Limbic Areas in Cortical Processing. Trends
1065 Cogn Sci *20*, 96-106. 10.1016/j.tics.2015.11.005.

1066 Chang, E.F., Raygor, K.P., and Berger, M.S. (2015). Contemporary model of language organization: an
1067 overview for neurosurgeons. J Neurosurg *122*, 250-261. 10.3171/2014.10.JNS132647.

1068 Cloutman, L.L. (2013). Interaction between dorsal and ventral processing streams: where, when and
1069 how? Brain Lang *127*, 251-263. 10.1016/j.bandl.2012.08.003.

1070 Cogan, G.B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., and Pesaran, B. (2014). Sensory-motor
1071 transformations for speech occur bilaterally. Nature *507*, 94-98. 10.1038/nature12935.

1072 Coifman, R.R., and Hirn, M.J. (2014). Diffusion maps for changing data. Applied and Computational
1073 Harmonic Analysis *36*, 79-107. https://doi.org/10.1016/j.acha.2013.03.001.

1074 Coifman, R.R., Lafon, S., Lee, A.B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S.W. (2005).
1075 Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps.
1076 Proc Natl Acad Sci U S A *102*, 7426-7431. 10.1073/pnas.0500334102.

1077 Craig, A.D. (2003). Interoception: the sense of the physiological condition of the body. Curr Opin
1078 Neurobiol *13*, 500-505. 10.1016/s0959-4388(03)00090-4.

1079 de Heer, W.A., Huth, A.G., Griffiths, T.L., Gallant, J.L., and Theunissen, F.E. (2017). The Hierarchical
1080 Cortical Organization of Human Speech Processing. J Neurosci *37*, 6539-6557. 10.1523/jneurosci.3267-
1081 16.2017.

1082 Deen, B., Koldewyn, K., Kanwisher, N., and Saxe, R. (2015). Functional Organization of Social Perception
1083 and Cognition in the Superior Temporal Sulcus. Cereb Cortex *25*, 4596-4609. 10.1093/cercor/bhv111.

1084 Destrieux, C., Fischl, B., Dale, A., and Halgren, E. (2010). Automatic parcellation of human cortical gyri
1085 and sulci using standard anatomical nomenclature. Neuroimage *53*, 1-15.
1086 10.1016/j.neuroimage.2010.06.010.

1087 Destrieux, C., Terrier, L.M., Andersson, F., Love, S.A., Cottier, J.P., Duvernoy, H., Velut, S., Janot, K., and
1088 Zemmoura, I. (2017). A practical guide for the identification of major sulcogyral structures of the human
1089 cortex. Brain Struct Funct *222*, 2001-2015. 10.1007/s00429-016-1320-z.

1090 Eisner, F., McGettigan, C., Faulkner, A., Rosen, S., and Scott, S.K. (2010). Inferior frontal gyrus activation
1091 predicts individual differences in perceptual learning of cochlear-implant simulations. J Neurosci *30*,
1092 7179-7186. 10.1523/JNEUROSCI.4040-09.2010.

1093 Feinsinger, A., Pouratian, N., Ebadi, H., Adolphs, R., Andersen, R., Beauchamp, M.S., Chang, E.F., Crone,
1094 N.E., Collinger, J.L., Fried, I., et al. (2022). Ethical commitments, principles, and practices guiding
1095 intracranial neuroscientific research in humans. Neuron *110*, 188-194. 10.1016/j.neuron.2021.11.011.

1096 Friederici, A.D. (2012). The cortical language circuit: from auditory perception to sentence
1097 comprehension. Trends Cogn Sci *16*, 262-268. 10.1016/j.tics.2012.04.001.

1098 Friederici, A.D., Meyer, M., and von Cramon, D.Y. (2000). Auditory language comprehension: an event-
1099 related fMRI study on the processing of syntactic and lexical information. Brain Lang *75*, 289-300.

1100 Fruhholz, S., Trost, W., and Kotz, S.A. (2016). The sound of emotions-Towards a unifying neural network
1101 perspective of affective sound processing. Neurosci Biobehav Rev *68*, 96-110.
1102 10.1016/j.neubiorev.2016.05.002.

1103 Geschwind, N. (1970). The organization of language and the brain. Science *170*, 940-944.
1104 10.1126/science.170.3961.940.

1105 Gonzalez Alam, T., McKeown, B.L.A., Gao, Z., Bernhardt, B., Vos de Wael, R., Margulies, D.S., Smallwood,
1106 J., and Jefferies, E. (2021). A tale of two gradients: differences between the left and right hemispheres
1107 predict semantic cognition. Brain Struct Funct. 10.1007/s00429-021-02374-w.

1108 Gorno-Tempini, M.L., Dronkers, N.F., Rankin, K.P., Ogar, J.M., Phengrasamy, L., Rosen, H.J., Johnson, J.K.,
1109 Weiner, M.W., and Miller, B.L. (2004). Cognition and anatomy in three variants of primary progressive
1110 aphasia. Ann Neurol *55*, 335-346. 10.1002/ana.10825.

1111 Hacker, C.D., Snyder, A.Z., Pahwa, M., Corbetta, M., and Leuthardt, E.C. (2017). Frequency-specific
1112 electrophysiologic correlates of resting state fMRI networks. Neuroimage *149*, 446-457.
1113 10.1016/j.neuroimage.2017.01.054.

1114 Hackett, T.A. (2015). Anatomic organization of the auditory cortex. Handb Clin Neurol *129*, 27-53.
1115 10.1016/B978-0-444-62630-1.00002-0.

1116 Hackett, T.A., Preuss, T.M., and Kaas, J.H. (2001). Architectonic identification of the core region in
1117 auditory cortex of macaques, chimpanzees, and humans. J Comp Neurol. *441*, 197-222.

1118 Hagoort, P. (2019). The neurobiology of language beyond single-word processing. Science *366*, 55-58.
1119 10.1126/science.aax0289.

1120 Hamilton, L.S., Oganian, Y., Hall, J., and Chang, E.F. (2021). Parallel and distributed encoding of speech
1121 across human auditory cortex. Cell *184*, 4626-4639 e4613. 10.1016/j.cell.2021.07.019.

1122 Hickok, G. (2009). The functional neuroanatomy of language. Phys Life Rev *6*, 121-143.
1123 10.1016/j.plrev.2009.06.001.

1124 Hickok, G. (2012). Computational neuroanatomy of speech production. Nat Rev Neurosci *13*, 135-145.
1125 10.1038/nrn3158.

1126 Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. Nat Rev Neurosci *8*,
1127 393-402. 10.1038/nrn2113.

1128 Hickok, G., and Poeppel, D. (2015). Neural basis of speech perception. Handb Clin Neurol *129*, 149-160.
1129 10.1016/B978-0-444-62630-1.00008-1.

1130 Hipp, J.F., Hawellek, D.J., Corbetta, M., Siegel, M., and Engel, A.K. (2012). Large-scale cortical correlation
1131 structure of spontaneous oscillatory activity. Nat Neurosci *15*, 884-890. 10.1038/nn.3101.

1132 Howard, M.A., Volkov, I.O., Mirsky, R., Garell, P.C., Noh, M.D., Granner, M., Damasio, H., Steinschneider,
1133 M., Reale, R.A., Hind, J.E., and Brugge, J.F. (2000). Auditory cortex on the human posterior superior
1134 temporal gyrus. J Comp Neurol. *416*, 79-92.

1135 Hu, B. (2003). Functional organization of lemniscal and nonlemniscal auditory thalamus. Exp Br Res *153*,
1136 543-549.

1137 Humphreys, G.F., Hoffman, P., Visser, M., Binney, R.J., and Lambon Ralph, M.A. (2015). Establishing task-
1138 and modality-dependent dissociations between the semantic and default mode networks. Proc Natl
1139 Acad Sci U S A *112*, 7857-7862. 10.1073/pnas.1422760112.

1140 Husain, F.T., and Schmidt, S.A. (2014). Using resting state functional connectivity to unravel networks of
1141 tinnitus. Hear Res *307*, 153-162. 10.1016/j.heares.2013.07.010.

1142 Jackson, R.L., Bajada, C.J., Rice, G.E., Cloutman, L.L., and Lambon Ralph, M.A. (2018). An emergent
1143 functional parcellation of the temporal cortex. Neuroimage *170*, 385-399.
1144 10.1016/j.neuroimage.2017.04.024.

1145 Jackson, R.L., Hoffman, P., Pobric, G., and Lambon Ralph, M.A. (2016). The Semantic Network at Work
1146 and Rest: Differential Connectivity of Anterior Temporal Lobe Subregions. J Neurosci *36*, 1490-1501.
1147 10.1523/JNEUROSCI.2999-15.2016.

1148 Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and
1149 accurate linear registration and motion correction of brain images. Neuroimage *17*, 825-841.

1150 Jones, E.G., Coulter, J.D., and Hendry, S.H. (1978). Intracortical connectivity of architectonic fields in the
1151 somatic sensory, motor and parietal cortex of monkeys. J Comp Neurol *181*, 291-347.
1152 10.1002/cne.901810206.

1153 Kaas, J.H., and Hackett, T.A. (1998). Subdivisions of auditory cortex and levels of processing in primates.
1154 Audiol Neurootol *3*, 73-85. 10.1159/000013783.

1155 Kaas, J.H., and Hackett, T.A. (2000). Subdivisions of auditory cortex and processing streams in primates.
1156 Proc Natl Acad Sci USA *97*, 11793-11799.

1157 Kahn, I., Andrews-Hanna, J.R., Vincent, J.L., Snyder, A.Z., and Buckner, R.L. (2008). Distinct cortical
1158 anatomy linked to subregions of the medial temporal lobe revealed by intrinsic functional connectivity. J
1159 Neurophysiol *100*, 129-139. 10.1152/jn.00077.2008.

1160 Karolis, V.R., Corbetta, M., and Thiebaut de Schotten, M. (2019). The architecture of functional
1161 lateralisation and its relationship to callosal connectivity in the human brain. Nat Commun *10*, 1417.
1162 10.1038/s41467-019-09344-1.

1163 Keitel, A., and Gross, J. (2016). Individual Human Brain Areas Can Be Identified from Their Characteristic
1164 Spectral Activation Fingerprints. PLoS Biol *14*, e1002498. 10.1371/journal.pbio.1002498.

1165 Kiebel, S.J., Daunizeau, J., and Friston, K.J. (2008). A hierarchy of time-scales and the brain. PLoS Comput
1166 Biol *4*, e1000209. 10.1371/journal.pcbi.1000209.

1167 Knecht, S., Drager, B., Deppe, M., Bobe, L., Lohmann, H., Floel, A., Ringelstein, E.B., and Henningsen, H.
1168 (2000). Handedness and hemispheric language dominance in healthy humans. Brain *123 Pt 12*, 2512-
1169 2518. 10.1093/brain/123.12.2512.

1170 Ko, A.L., Weaver, K.E., Hakimian, S., and Ojemann, J.G. (2013). Identifying functional networks using
1171 endogenous connectivity in gamma band electrocorticography. Brain Connect *3*, 491-502.
1172 10.1089/brain.2013.0157.

1173 Kovach, C.K., and Gander, P.E. (2016). The demodulated band transform. J Neurosci Methods *261*, 135-
1174 154. 10.1016/j.jneumeth.2015.12.004.

1175 Kraus, K.S., and Canlon, B. (2012). Neuronal connectivity and interactions between the auditory and
1176 limbic systems. Effects of noise and tinnitus. Hear Res *288*, 34-46. 10.1016/j.heares.2012.02.009.

1177 Kuehn, E., Mueller, K., Lohmann, G., and Schuetz-Bosbach, S. (2016). Interoceptive awareness changes
1178 the posterior insula functional connectivity profile. Brain Struct Funct *221*, 1555-1571. 10.1007/s00429-
1179 015-0989-8.

1180 Kumar, S., Gander, P.E., Berger, J.I., Billig, A.J., Nourski, K.V., Oya, H., Kawasaki, H., Howard, M.A., 3rd,
1181 and Griffiths, T.D. (2021). Oscillatory correlates of auditory working memory examined with human
1182 electrocorticography. Neuropsychologia *150*, 107691. 10.1016/j.neuropsychologia.2020.107691.

1183 Kumar, S., Joseph, S., Gander, P.E., Barascud, N., Halpern, A.R., and Griffiths, T.D. (2016). A Brain System
1184 for Auditory Working Memory. J Neurosci *36*, 4492-4505. 10.1523/jneurosci.4341-14.2016.

1185 Lambon Ralph, M.A., Jefferies, E., Patterson, K., and Rogers, T.T. (2017). The neural and computational
1186 bases of semantic cognition. Nat Rev Neurosci *18*, 42-55. 10.1038/nrn.2016.150.

1187 Leaver, A.M., and Rauschecker, J.P. (2010). Cortical representation of natural complex sounds: effects of
1188 acoustic features and auditory object category. J Neurosci *30*, 7604-7612. 10.1523/jneurosci.0296-
1189 10.2010.

1190 Makris, N., Papadimitriou, G.M., Kaiser, J.R., Sorg, S., Kennedy, D.N., and Pandya, D.N. (2009).
1191 Delineation of the middle longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. Cereb
1192 Cortex *19*, 777-785. 10.1093/cercor/bhn124.

1193 Maller, J.J., Welton, T., Middione, M., Callaghan, F.M., Rosenfeld, J.V., and Grieve, S.M. (2019). Revealing
1194 the Hippocampal Connectome through Super-Resolution 1150-Direction Diffusion MRI. Sci Rep *9*, 2418.
1195 10.1038/s41598-018-37905-9.

1196 Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G., Bezgin, G., Eickhoff,
1197 S.B., Castellanos, F.X., Petrides, M., et al. (2016). Situating the default-mode network along a principal

1198 gradient of macroscale cortical organization. Proc Natl Acad Sci U S A *113*, 12574-12579.
1199 10.1073/pnas.1608282113.

1200 McGettigan, C., and Scott, S.K. (2012). Cortical asymmetries in speech perception: what's wrong, what's
1201 right and what's left? Trends Cogn Sci *16*, 269-276. 10.1016/j.tics.2012.04.006.

1202 Mesulam, M.M. (1998). From sensation to cognition. Brain *121 (Pt 6)*, 1013-1052.

1203 Mesulam, M.M. (2000). Paralimbic (mesocortical) areas. In Principles of behavioral and cognitive
1204 neurology, (Oxford University Press), pp. 49-54.

1205 Michelmann, S., Price, A.R., Aubrey, B., Strauss, C.K., Doyle, W.K., Friedman, D., Dugan, P.C., Devinsky,
1206 O., Devore, S., Flinker, A., et al. (2021). Moment-by-moment tracking of naturalistic learning and its
1207 underlying hippocampo-cortical interactions. Nat Commun *12*, 5394. 10.1038/s41467-021-25376-y.

1208 Moerel, M., De Martino, F., and Formisano, E. (2014). An anatomical and functional topography of
1209 human auditory cortical areas. Front Neurosci *8*, 225. 10.3389/fnins.2014.00225.

1210 Morel, A., Garraghty, P.E., and Kaas, J.H. (1993). Tonotopic organization, architectonic fields, and
1211 connections of auditory cortex in macaque monkeys. J Comp Neurol *335*, 437-459.
1212 10.1002/cne.903350312.

1213 Munoz-Lopez, M., Insausti, R., Mohedano-Moriano, A., Mishkin, M., and Saunders, R.C. (2015).
1214 Anatomical pathways for auditory memory II: information from rostral superior temporal gyrus to
1215 dorsolateral temporal pole and medial temporal cortex. Front Neurosci *9*, 158.
1216 10.3389/fnins.2015.00158.

1217 Munoz-Lopez, M.M., Mohedano-Moriano, A., and Insausti, R. (2010). Anatomical pathways for auditory
1218 memory in primates. Front Neuroanat *4*, 129. 10.3389/fnana.2010.00129.

1219 Nourski, K.V., and Howard, M.A., 3rd (2015). Invasive recordings in the human auditory cortex. Handb
1220 Clin Neurol *129*, 225-244. 10.1016/B978-0-444-62630-1.00013-5.

1221 Nourski, K.V., Steinschneider, M., Oya, H., Kawasaki, H., Jones, R.D., and Howard, M.A. (2014). Spectral
1222 organization of the human lateral superior temporal gyrus revealed by intracranial recordings. Cereb
1223 Cortex *24*, 340-352. 10.1093/cercor/bhs314.

1224 Nourski, K.V., Steinschneider, M., and Rhone, A.E. (2016). Electrocorticographic Activation within Human
1225 Auditory Cortex during Dialog-Based Language and Cognitive Testing. Front Hum Neurosci *10*, 202.
1226 10.3389/fnhum.2016.00202.

1227 Nourski, K.V., Steinschneider, M., Rhone, A.E., and Howard Iii, M.A. (2017). Intracranial
1228 Electrophysiology of Auditory Selective Attention Associated with Speech Classification Tasks. Front
1229 Hum Neurosci *10*, 691. 10.3389/fnhum.2016.00691.

1230 Nourski, K.V., Steinschneider, M., Rhone, A.E., Kovach, C.K., Banks, M.I., Krause, B.M., Kawasaki, H., and
1231 Howard, M.A. (2021). Electrophysiology of the Human Superior Temporal Sulcus during Speech
1232 Processing. Cereb Cortex *31*, 1131-1148. 10.1093/cercor/bhaa281.

Nourski, K.V., Steinschneider, M., Rhone, A.E., Kovach, C.K., Kawasaki, H., and Howard, M.A., 3rd (2022). Gamma Activation and Alpha Suppression within Human Auditory Cortex during a Speech Classification Task. J Neurosci *42*, 5034-5046. 10.1523/JNEUROSCI.2187-21.2022.

Olson, I.R., McCoy, D., Klobusicky, E., and Ross, L.A. (2013). Social cognition and the anterior temporal lobes: a review and theoretical framework. Soc Cogn Affect Neurosci *8*, 123-133. 10.1093/scan/nss119.

Patterson, K., Nestor, P.J., and Rogers, T.T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci *8*, 976-987. 10.1038/nrn2277.

Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage *59*, 2142-2154. 10.1016/j.neuroimage.2011.10.018.

Price, C.J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage *62*, 816-847. 10.1016/j.neuroimage.2012.04.062.

Rauschecker, J.P. (2018). Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. Cortex *98*, 262-268. 10.1016/j.cortex.2017.10.020.

Rauschecker, J.P., and Scott, S.K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci *12*, 718-724. 10.1038/nn.2331.

Remedios, R., Logothetis, N.K., and Kayser, C. (2009). An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. J Neurosci *29*, 1034-1045. 10.1523/JNEUROSCI.4089-08.2009.

Rocchi, F., Oya, H., Balezeau, F., Billig, A.J., Kocsis, Z., Jenison, R.L., Nourski, K.V., Kovach, C.K., Steinschneider, M., Kikuchi, Y., et al. (2021). Common fronto-temporal effective connectivity in humans and monkeys. Neuron *109*, 852-868 e858. 10.1016/j.neuron.2020.12.026.

Rohr, K., Stiehl, H.S., Sprengel, R., Buzug, T.M., Weese, J., and Kuhn, M.H. (2001). Landmark-based elastic registration using approximating thin-plate splines. IEEE Trans Med Imaging *20*, 526-534. 10.1109/42.929618.

Rolls, E.T. (2019). The cingulate cortex and limbic systems for emotion, action, and memory. Brain Struct Funct *224*, 3001-3018. 10.1007/s00429-019-01945-2.

Satopaa, V., Albrecht, J., Irwin, D., and Raghavan, B. (2011). Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior. 20-24 June 2011. pp. 166-171.

Saur, D., Kreher, B.W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M.S., Umarova, R., Musso, M., Glauche, V., Abel, S., et al. (2008). Ventral and dorsal pathways for language. Proc Natl Acad Sci U S A *105*, 18035-18040. 10.1073/pnas.0805234105.

Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.N., Holmes, A.J., Eickhoff, S.B., and Yeo, B.T.T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cereb Cortex *28*, 3095-3114. 10.1093/cercor/bhx179.

1269 Schirmer, A., Fox, P.M., and Grandjean, D. (2012). On the spatial organization of sound processing in the
1270 human temporal lobe: a meta-analysis. Neuroimage *63*, 137-147. 10.1016/j.neuroimage.2012.06.025.

1271 Scott, S.K. (2012). The neurobiology of speech perception and production--can functional imaging tell us
1272 anything we did not already know? J Commun Disord *45*, 419-425. 10.1016/j.jcomdis.2012.06.007.

1273 Scott, S.K., Blank, C.C., Rosen, S., and Wise, R.J. (2000). Identification of a pathway for intelligible speech
1274 in the left temporal lobe. Brain *123 Pt 12*, 2400-2406. 10.1093/brain/123.12.2400.

1275 Seitzman, B.A., Snyder, A.Z., Leuthardt, E.C., and Shimony, J.S. (2019). The State of Resting State
1276 Networks. Top Magn Reson Imaging *28*, 189-196. 10.1097/RMR.0000000000000214.

1277 Sherman, S.M., and Guillery, R.W. (2011). Distinct functions for direct and transthalamic corticocortical
1278 connections. J Neurophysiol *106*, 1068-1077. jn.00429.2011 [pii];10.1152/jn.00429.2011 [doi].

1279 Simmons, W.K., and Martin, A. (2009). The anterior temporal lobes and the functional architecture of
1280 semantic memory. J Int Neuropsychol Soc *15*, 645-649. 10.1017/S1355617709990348.

1281 Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro,
1282 R., Laird, A.R., and Beckmann, C.F. (2009). Correspondence of the brain's functional architecture during
1283 activation and rest. Proc Natl Acad Sci U S A *106*, 13040-13045. 10.1073/pnas.0905267106.

1284 Spitsyna, G., Warren, J.E., Scott, S.K., Turkheimer, F.E., and Wise, R.J. (2006). Converging language
1285 streams in the human temporal lobe. J Neurosci *26*, 7328-7336. 10.1523/JNEUROSCI.0559-06.2006.

1286 Steinschneider, M., Nourski, K.V., and Fishman, Y.I. (2013). Representation of speech in human auditory
1287 cortex: is it special? Hear Res *305*, 57-73. 10.1016/j.heares.2013.05.013.

1288 Steinschneider, M., Nourski, K.V., Rhone, A.E., Kawasaki, H., Oya, H., and Howard, M.A., 3rd (2014).
1289 Differential activation of human core, non-core and auditory-related cortex during speech categorization
1290 tasks as revealed by intracranial recordings. Front Neurosci *8*, 240. 10.3389/fnins.2014.00240.

1291 Turkeltaub, P.E., and Coslett, H.B. (2010). Localization of sublexical speech perception components.
1292 Brain Lang *114*, 1-15. 10.1016/j.bandl.2010.03.008.

1293 Upadhyay, J., Silver, A., Knaus, T.A., Lindgren, K.A., Ducros, M., Kim, D.S., and Tager-Flusberg, H. (2008).
1294 Effective and structural connectivity in the human auditory cortex. J Neurosci *28*, 3341-3349.
1295 10.1523/JNEUROSCI.4434-07.2008.

1296 van den Heuvel, M.P., and Sporns, O. (2013). Network hubs in the human brain. Trends Cogn Sci *17*, 683-
1297 696. 10.1016/j.tics.2013.09.012.

1298 Venezia, J.H., Vaden, K.I., Jr., Rong, F., Maddox, D., Saberi, K., and Hickok, G. (2017). Auditory, Visual and
1299 Audiovisual Speech Processing Streams in Superior Temporal Sulcus. Front Hum Neurosci *11*, 174.
1300 10.3389/fnhum.2017.00174.

1301 Visser, M., Jefferies, E., and Lambon Ralph, M.A. (2010). Semantic processing in the anterior temporal
1302 lobes: a meta-analysis of the functional neuroimaging literature. J Cogn Neurosci *22*, 1083-1094.
1303 10.1162/jocn.2009.21309.

1304 Wang, J., Tao, A., Anderson, W.S., Madsen, J.R., and Kreiman, G. (2021). Mesoscopic physiological
1305 interactions in the human brain reveal small-world properties. Cell Rep *36*, 109585.
1306 10.1016/j.celrep.2021.109585.

1307 Wang, S.F., Ritchey, M., Libby, L.A., and Ranganath, C. (2016). Functional connectivity based parcellation
1308 of the human medial temporal lobe. Neurobiol Learn Mem *134 Pt A*, 123-134.
1309 10.1016/j.nlm.2016.01.005.

1310 Wilson, S.M., Bautista, A., and McCarron, A. (2018). Convergence of spoken and written language
1311 processing in the superior temporal sulcus. Neuroimage *171*, 62-74. 10.1016/j.neuroimage.2017.12.068.

1312 Woods, D.L., and Alain, C. (2009). Functional imaging of human auditory cortex. Curr Opin Otolaryngol
1313 Head Neck Surg *17*, 407-411. 10.1097/MOO.0b013e3283303330.

1314 Woods, D.L., Herron, T.J., Cate, A.D., Yund, E.W., Stecker, G.C., Rinne, T., and Kang, X. (2010). Functional
1315 properties of human auditory cortical fields. Front Syst Neurosci *4*, 155. 10.3389/fnsys.2010.00155.

1316 Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller,
1317 J.W., Zollei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by
1318 intrinsic functional connectivity. J Neurophysiol *106*, 1125-1165. 10.1152/jn.00338.2011.

1319 Zachlod, D., Rüttgers, B., Bludau, S., Mohlberg, H., Langner, R., Zilles, K., and Amunts, K. (2020). Four
1320 new cytoarchitectonic areas surrounding the primary and early auditory cortex in human brains. Cortex
1321 *128*, 1-21. 10.1016/j.cortex.2020.02.021.

1322 Zhang, Y., Ding, Y., Huang, J., Zhou, W., Ling, Z., Hong, B., and Wang, X. (2021). Hierarchical cortical
1323 networks of "voice patches" for processing voices in human brain. Proc Natl Acad Sci U S A *118*.
1324 10.1073/pnas.2113887118.

1325 Zhang, Y., Zhou, W., Wang, S., Zhou, Q., Wang, H., Zhang, B., Huang, J., Hong, B., and Wang, X. (2019).
1326 The Roles of Subdivisions of Human Insula in Emotion Perception and Auditory Processing. Cereb Cortex
1327 *29*, 517-528. 10.1093/cercor/bhx334.

1328