


## RESEARCH ARTICLE

# Diffusion magnetic resonance imaging connectome features are predictive of functional lateralization of semantic processing in the anterior temporal lobes

Felix Zahnert<sup>1</sup>  | Gunter Kräling<sup>2</sup> | Leander Melms<sup>3</sup> | Marcus Belke<sup>1,4</sup> | Urs Kleinholdermann<sup>5</sup> | Lars Timmermann<sup>5,6,7</sup> | Martin Hirsch<sup>3</sup> | Andreas Jansen<sup>6,7,8</sup> | Peter Mross<sup>1</sup> | Katja Menzler<sup>1,6,7</sup> | Lena Habermehl<sup>1</sup> | Susanne Knake<sup>1,4,6,7</sup>

<sup>1</sup>Epilepsy Center Hesse, Department for Neurology, University Hospital Marburg, Philipps University Marburg, Marburg, Germany

<sup>2</sup>Department of Medical Technology, University Hospital Marburg, Marburg, Germany

<sup>3</sup>Institute for Artificial Intelligence, University Hospital Marburg, Philipps University Marburg, Marburg, Germany

<sup>4</sup>LOEWE Center for Personalized Translational Epilepsy Research (CePTER), Goethe-University Frankfurt, Frankfurt Am Main, Germany

<sup>5</sup>Department for Neurology, University Hospital Marburg, Philipps University Marburg, Marburg, Germany

<sup>6</sup>Center for Mind, Brain and Behavior (CMBB), Philipps-University Marburg, Marburg, Germany

<sup>7</sup>Core Facility Brainimaging, Faculty of Medicine, University of Marburg, Marburg, Germany

<sup>8</sup>Department for Psychiatry and Psychotherapy, University Hospital Marburg, Philipps University Marburg, Marburg, Germany

## Correspondence

Felix Zahnert, Epilepsy Center Hesse, Department for Neurology, University Hospital Marburg, Philipps University Marburg, 35043 Marburg, Germany.  
Email: [zahnert@med.uni-marburg.de](mailto:zahnert@med.uni-marburg.de)

## Funding information

McDonnell Center for Systems Neuroscience at Washington University; NIH Blueprint for

## Abstract

Assessment of regional language lateralization is crucial in many scenarios, but not all populations are suited for its evaluation via task-functional magnetic resonance imaging (fMRI). In this study, the utility of structural connectome features for the classification of language lateralization in the anterior temporal lobes (ATLs) was investigated. Laterality indices for semantic processing in the ATL were computed from task-fMRI in 1038 subjects from the Human Connectome Project who were labeled as stronger rightward lateralized (RL) or stronger leftward to bilaterally lateralized (LL) in a data-driven approach. Data of unrelated subjects ( $n = 432$ ) were used for further analyses. Structural connectomes were generated from diffusion-MRI tractography, and graph theoretical metrics (node degree, betweenness centrality) were computed. A neural network (NN) and a random forest (RF) classifier were trained on these metrics to classify subjects as RL or LL. After classification, comparisons of network measures were conducted via permutation testing. Degree-based classifiers produced significant above-chance predictions both during cross-validation (NN: AUC-ROC[CI] = 0.68[0.64–0.73], accuracy[CI] = 68.34%[63–73.2%]; RF: AUC-ROC[CI] = 0.7[0.66–0.73], accuracy[CI] = 64.81%[60.9–68.5]) and testing (NN: AUC-ROC[CI] = 0.69[0.53–0.84], accuracy[CI] = 68.09[53.2–80.9]; RF: AUC-ROC[CI] = 0.68[0.53–0.84], accuracy[CI] = 68.09[55.3–80.9]). Comparison of network metrics revealed small effects of increased node degree within the right posterior middle temporal gyrus (pMTG) in subjects with RL, while degree was decreased in the right posterior cingulate cortex (PCC). Above-chance predictions of functional language lateralization in the ATL are possible based on diffusion-MRI connectomes alone. Increased degree within the right pMTG as a right-sided homologue of a known semantic hub, and decreased hubness of the right PCC may form a structural basis for rightward-lateralized semantic processing.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

## KEYWORDS

graph theory, language dominance, machine learning, semantic processing, structural connectome

## 1 | INTRODUCTION

It is well established that the language system is lateralized to the left hemisphere in the majority of individuals, while a minority shows a rightward lateralization (RL; Knecht, Deppe, et al., 2000; Knecht, Dräger, et al., 2000). This lateralization is commonly assessed via language task-functional magnetic resonance imaging (tfMRI). Word fluency tasks, for example, produce strongly lateralized activations especially in language regions within the frontal lobe (Bradshaw et al., 2017), and tfMRI can lead to robust estimates of hemispheric language dominance in these regions of interest (ROIs) (Fernández et al., 2003; Jansen et al., 2006). It has, however, been shown that language function is distributed over a large, partially bihemispheric network (Binder et al., 2009; Branco et al., 2020; Price, 2012), whose regional involvement can differ dramatically across tasks and individuals (Seghier, 2008). This implies a potential benefit of tfMRI protocols tailored to fit the ROIs under investigation, which can be relevant, for example, in clinical scenarios such as prior to epilepsy surgery (Tailby et al., 2017).

Binder et al. proposed an overt semantic processing task that robustly activates the anterior temporal lobes (ATLs) with a bilateral, on average left dominant pattern (Binder et al., 2011). It has been demonstrated that the ATLs are necessary for lateralized spoken word processing in patients with semantic dementia (Cope et al., 2020), and that the resection of the language-dominant ATL can lead to deficits in naming and semantic processing (Sabsevitz et al., 2003). Therefore, this task is a promising candidate for assessing regional lateralization of semantic processing within the ATL.

However, many subjects, for example, in clinical populations, are unable to adequately perform tfMRI. Therefore, task free alternatives for delineating regional language dominance are of great interest. There are several approaches to derive lateralization indices from other modalities. Language networks and individual activations from language tfMRI have been successfully predicted from resting-state functional connectivity (rs-fc; Parker Jones et al., 2017; Tavor et al., 2016; Tomasi & Volkow, 2012) and laterality indices (LIs) were derived from these with varying degrees of correlation to tfMRI-LI in healthy (Tavor et al., 2016) and clinical (Desai et al., 2018; Doucet et al., 2015) populations. In the latter, rs-fMRI-based classification of language dominance was found to show up to 63% concordance with the tfMRI-based classification (Rolinski et al., 2020). Others have demonstrated correlations of word-generation-task-based tfMRI-LI with rs-fc of the left inferior frontal gyrus (IFG) and with global graph theoretical (GT) metrics obtained from rs-fc (Wang et al., 2019).

There have also been efforts to infer language lateralization from diffusion MRI (dMRI). Tractometry of the arcuate fasciculus (AF) has commonly revealed a leftward lateralization in volume, fractional

anisotropy (FA) and fiber count (Delgado-Fernández et al., 2020; Matsumoto et al., 2008; Silva & Citterio, 2017; Tiwari et al., 2011; Vernooij et al., 2007). One study of 10 right-handed healthy participants extended on this knowledge and the authors detected correlations of LI obtained from task-fMRI with LI obtained from FA within bilateral fronto-temporal tracts between language regions (verb generation:  $r = .782$ , reading comprehension:  $r = .651$ ) (Powell et al., 2006). Recently, it has been demonstrated that the lateralization of the AF, however, does not discriminate between subjects with right- and left-sided functional language dominance (Verhelst et al., 2021). Another study demonstrated a significant negative correlation of the LI of FA within the direct segments of the AF with interhemispheric rs-fc of bilateral Broca's territories ( $r = -.65$ ), while again, the LI of the AF was not correlated with task-evoked asymmetry or asymmetry of intrahemispheric rs-fc (Piervincenzi et al., 2016). According to the authors, this might potentially signify a downregulation of interhemispheric connectivity (via the corpus callosum) mediated by this leftward structural asymmetry. Interestingly, small effects of increased FA within the corpus callosum have been associated with atypical language dominance in both healthy subjects ( $R^2 = .082$ ) (Häberling et al., 2011) and in patients with arteriovenous malformations (Li et al., 2021). One study has demonstrated a correlation between LI derived from tfMRI and an asymmetry of Meyer's loop in patients with epilepsy (Nowell et al., 2016).

Overall, tract-based and microstructural properties have commonly been assessed in subjects with left and more rightward language lateralization. However, as opposed to rs-fc, the role of the dMRI connectome in subjects with differential hemispheric language lateralization is unclear.

The connectome is a description of brain connectivity and can be modeled by brain graphs of interregional connections, which can be estimated with different methodologies such as dMRI tractography (Bullmore & Bassett, 2011; Fornito & Bullmore, 2015). Graph theory is a branch of mathematics dedicated to the analysis of networks (Bassett & Sporns, 2017). In a brain network, its elements (e.g., brain regions) are termed nodes, and their connections or relationships are termed edges (Sporns, 2014).

Structural connectomes are brain networks generated from dMRI tractography and constitute graphs of white matter interconnections of its nodes (Sotiropoulos & Zalesky, 2019). Based on such networks, various graph-theoretical metrics can be computed that describe topological properties of the network (Sporns, 2014). One such metric, for example, is node degree, which simply describes the number of connections that link a node to the rest of the network (Bullmore & Sporns, 2009). GT properties have been investigated in a wide range of fields to characterize networks in healthy subjects and clinical populations (Bullmore & Sporns, 2009; Sporns, 2014).

GT measures derived from resting-state fMRI connectomes, for example, have been successfully applied to predict naming decline after ATL surgery (Audrain et al., 2018). In terms of machine learning, GT metrics seem attractive, as they summarize properties of a network and thus can lead to a reduction of input features for model training.

The present study aimed to investigate the utility of various structural connectome features for predicting the regional lateralization of semantic processing within the ATL. In a second step, we sought to delineate a structural basis of functional lateralization. To this end, the large sample of high-quality neuroimaging data from the Human Connectome Project (HCP; van Essen et al., 2013), including the above-mentioned semantic processing task was leveraged (Binder et al., 2011).

Subjects were grouped into rightward and more leftward lateralized subjects and machine learning models were trained on structural connectome features to differentiate between both groups. We then explored whether differences in nodal hubness are present in subjects with differentially lateralized semantic processing.

## 2 | METHODS

### 2.1 | Subjects

All subjects from the HCP S1200 release in whom a minimally processed structural MRI (Freesurfer-Pipeline), language task-fMRI and diffusion weighted MRI (BedpostX-processed) were available ( $n = 1039$ ) (van Essen et al., 2013) were included. One subject was excluded due to lack of any ATL tfMRI-activation. Subjects were healthy adults aged 22–35 years, 558 (53.8%) were female and 46.2% were male. After estimation of a threshold for rightward language laterality (see Section 2.4.1), a subset of 432 unrelated subjects was selected for further analyses. To this end, 117 unrelated subjects with RL were identified, and all subjects with LL who were neither related to the subjects with RL, nor to each other, were included ( $n = 315$ ). Analyses including all 1038 subjects are reported in the supplementary material.

### 2.2 | MRI acquisition and preprocessing

The HCP acquisition protocols and preprocessing pipelines are described in detail elsewhere (Glasser et al., 2013; Sotiropoulos et al., 2013; Uğurbil et al., 2013). In short, dMRI was acquired at a resolution of  $1.25 \text{ mm}^3$  with 90 gradient directions at each of three shells with  $b$ -values of 1000, 2000, and  $3000 \text{ s/m}^2$ , respectively. Images were corrected for eddy current distortions, movement and susceptibility induced distortions. Images were registered to native space and fiber orientations had been estimated using FSL's BedpostX (Jbabdi et al., 2012). Task-fMRI was acquired with a gradient-echo EPI sequence (TE 33.1 ms, TR 720 ms, multiband factor = 8) over two runs with left–right and right–left phase encoding respectively. Data

from the fMRI volume preprocessing pipeline was used, to which gradient distortion correction, motion correction, TOPUP, bias field removal as well as registration to the native T1 weighted image and to standard space had been applied. Structural scans (3D MPRAGE, TE 2.14 ms, TR 2400 ms, 0.7 mm isotropic) had already been processed via a modified version of Freesurfer's recon-all pipeline.

### 2.3 | fMRI language task

The language task implemented in the HCP task fMRI battery is a semantic processing task first proposed and described in detail by Binder et al. (2011). Subjects completed 26 runs each of a semantic processing task (“Story”) and a math task as control.

During the story task, participants were presented short stories from Aesop's fables (five to nine sentences), after each of which they were asked dichotomous questions about the story's content. The choice was given by the participant via right-hand button presses. During the control task, participants were presented math problems in varying and adjustable degree of difficulty, again with a dichotomous possibility to answer (Binder et al., 2011).

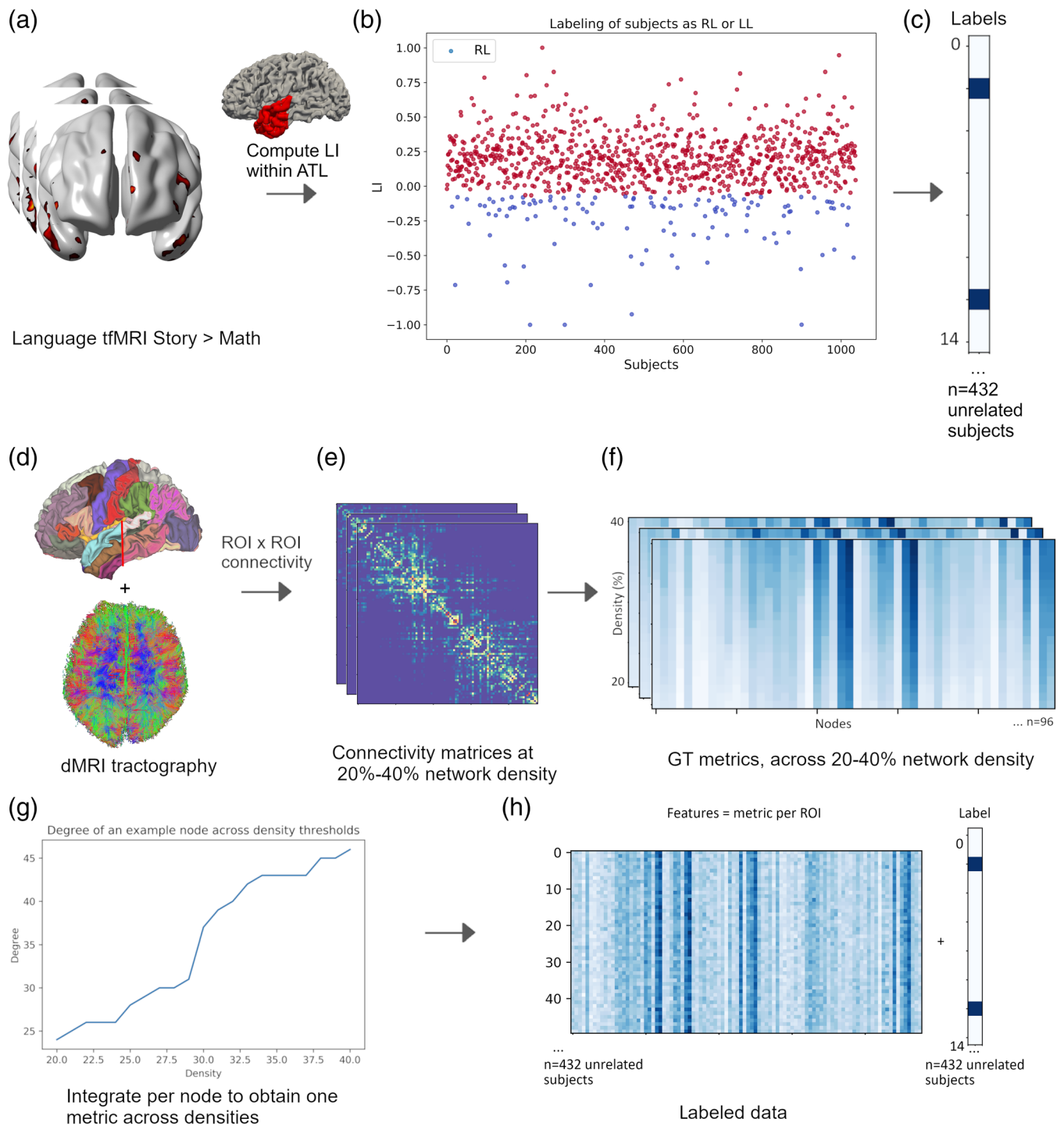
### 2.4 | MRI processing

#### 2.4.1 | Functional MRI processing and calculation of LIs

FSL Feat (FSL 6.0) was used for fMRI analysis (Woolrich et al., 2001). As we were interested in the lateralization of semantic processing, only the Story > Math contrast was calculated. In each subject, one acquisition with right to left- and one run with left to right phase encoding was present. First-level analysis was carried out on each of these images without application of any thresholding, using a double-gamma-HRF convolution. Subject-wise second level analysis was performed to create an average of these images and cluster-wise thresholding of activations was applied at a cluster-forming threshold of  $Z = 3.1$  and a cluster  $p$ -threshold of .05 (Figure 1a). An ATL ROI anterior to MNI  $Y = -22$  (Murphy et al., 2017) was generated in standard MNI space (Figure 1a).

Right and left ATL activations were calculated by summing the  $z$ -statistics of each significantly activated voxel in each ATL ROI separately, resulting in a weighted measure of task-activation per ROI. LIs were calculated as follows:  $I = \frac{\text{Left} - \text{Right}}{\text{Left} + \text{Right}}$ , so that LI ranged from  $-1$  to  $1$ . Generally, LIs can be calculated based on the extent or the magnitude of activations in the respective ROI, and it has been demonstrated that magnitude-based calculations can yield more robust estimates of regional language dominance (Jansen et al., 2006). Other approaches that avoid a fixed activation threshold above which LI are calculated are available as well (Matsuo et al., 2012; Wilke & Lidzba, 2007).

For binary classification of language lateralization, a threshold needed to be determined below which an LI was considered more



**FIGURE 1** Labeling and preprocessing of connectomes. (a) Extraction of labels for the data. Laterality indices were calculated from language-task-functional magnetic resonance imaging (tfMRI) using the Story > Math contrast within the anterior temporal lobe (ATL). (b) Distribution of laterality indices and assignment of labels (see Section 2.4.1). Subjects (x axis) in blue were defined as more rightward lateralized. (c) Binary label as RL or LL per subject after selection of 432 unrelated subjects. (d) Tractography with computation of regions of interest (ROI) × ROI connectivity between parcels of the modified Desikan parcellation as detailed in Section 2.4.2. The red line within the temporal lobe highlights the coordinate MNI  $Y = -22$ , at which temporal lobe seeds were split into anterior and posterior divisions. (e) Resulting preprocessed connectivity matrices, thresholded at network densities of 20–40%. The depicted example networks had a density of 30%. (f) Graph theoretical metrics were computed in each subject and at each density. (g) A curve of the GT metric as a function of network density resulted. Values under curve were averaged to obtain one metric across network densities. (h) For each subject and metric, a vector of 96 features (corresponding to 96 ROIs) as well as the respective label of functional lateralization resulted. Each metric was used separately for model training. GT, graph theory; LI, laterality index; LL, leftward and bilateral lateralization; RL, stronger rightward lateralization

strongly rightward lateralized for this task. In the literature, leftward lateralization of language function is commonly considered as typical and it is often assumed at a LI of  $>0.2$ – $0.25$  (Janecek et al., 2013; Seghier, 2008). An  $|LI| < 0.2$  usually is considered as bilateral, and  $LI < -0.2$  are considered as rightward lateralized (Jansen et al., 2006; Seghier, 2008). However, these thresholds have commonly been set for word fluency tasks, and the degree to which different tasks lateralize language function can vary (Seghier, 2008).

The present task has been reported to typically produce bilateral, on average left dominant activations (Barch et al., 2013; Binder et al., 2011). We computed the distribution of left, bilateral, and right lateralized subjects according to canonical thresholds of  $LI = |0.2|$  and confirmed these results (mean  $LI = 0.16$ ,  $SD = 0.23$ ; left: 446 subjects (43%), bilateral: 542 (52.2%), right: 50 (5%)). According to these thresholds of LI in the case of this only moderately lateralizing task, an implausible proportion of subjects would have had to be deemed “atypical” (57%).

Thus, a cutoff for stronger RL lateralization was identified in a data-driven manner based on the distribution of LI among all subjects. Data-driven classification of language lateralization has been conducted in previous literature, albeit less frequently than adoption of the canonical thresholds (Seghier, 2008). In this study, this data-driven approach allowed for functional classification of regional laterality as stronger left to bilateral lateralization (from here: LL) or RL based on this moderately lateralizing task.

The threshold for RL was computed as follows:  $mean_{LI} - SD_{LI} = -0.06805$ , and subjects with LI below this value were labeled as RL. Note that the above approach was nonetheless still an approximation to the unknown ground truth as to which range of LI truly is “typical” for this task. Therefore, classifications and statistics were also conducted using a different approach of threshold selection (k-means clustering, obtaining a threshold for RL at  $LI = -0.064318$ ), and these results are reported in the supplementary material.

After generating two subpopulations according to LI and selection of the subgroup of 432 unrelated participants, third level analysis was performed within FSL to compare task-activations in subjects with RL versus LL using a mixed effects model with a cluster-forming threshold of  $Z = 3.1$  and a cluster  $p$ -threshold of .05. This was done to infer which ATL-subregions, on average, drove lateralization and to gain insights on which other regions showed significant between-group differences in concert with lateralized ATL-activation.

## 2.4.2 | dMRI and tractography

The GPU Version of ProbtrackX2 (FSL) (Hernandez-Fernandez et al., 2019) was used for probabilistic tractography in all subjects. The Desikan parcellation was modified and chosen for seed generation (Desikan et al., 2006), including supratentorial subcortical ROI. Temporal lobe seeds were split into anterior and posterior divisions using a custom generated ATL mask (anterior MNI  $Y = -22$ ) in each individual. Cortical seeds were confined to voxels at the gray-/white-matter interface and the residual cortical gray matter was excluded

from tractography to avoid spurious connections across adjacent gyri. This way, 96 cortical and subcortical seeds were created for tractography and to estimate interregional connectivity.

To evaluate robustness of our results across parcellations, structural connectomes were additionally generated using the Schaefer 200 ROI functional parcellation (Schaefer et al., 2018) including 14 additional subcortical ROI. Results obtained from these connectomes are reported in the supplementary material.

Cerebrospinal fluid, ventricles, and infratentorial structures were excluded from the analysis and all white matter was used as a way-point mask. Five thousand streamlines were generated per voxel in each seed-ROI at the white–gray matter boundary and correction for streamline-length was applied.

## 2.5 | Preprocessing of network matrices

Tractography (Figure 1d) resulted in an asymmetric  $96 \times 96$  (for 96 ROIs) matrix  $A$  per subject, where each element  $A_{ij}$  indicated the streamline-count between two regions  $i$  and  $j$ . We applied a form of fractional scaling by dividing each element  $A_{ij}$  in a row (corresponding to a region  $i$ ) by the respective sum of all streamlines of that ROI that reached a target:

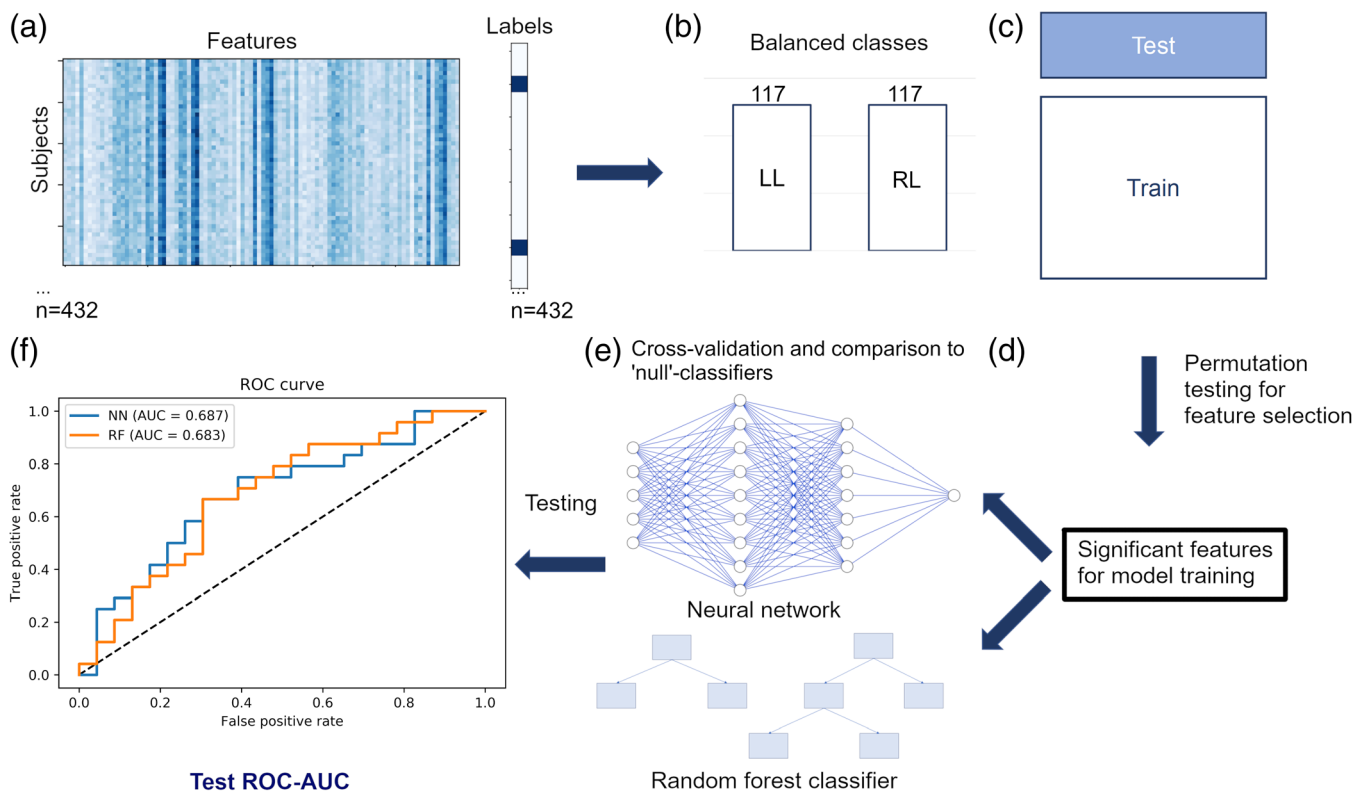
$$A_{ij, \text{scaled}} = \frac{A_{ij}}{\sum_{j=1}^N A_{ij}}$$

It has been demonstrated that fractional scaling of matrices derived from probabilistic tractography can lead to improved contrast between edges (Sotiropoulos & Zalesky, 2019). Scaled streamline counts were used as connectivity weights. Connectivity matrices were symmetrized elementwise according to  $\frac{A_{ij} + A_{ji}}{2}$ , except for connections where at least one of the edges was 0, as here both edges were set to 0 (Figure 1d). In the next step, connectivity matrices were thresholded to account for potential spurious connections inherent to probabilistic tractography (Sarwar et al., 2019).

The `get_components` function from the Brain Connectivity Toolbox for Python (<https://pypi.org/project/bctpy/>, (Rubinov & Sporns, 2010)) was used to assess whether any of the created graphs fragmented at a given density threshold (10, 15, 20, 30, and 40% connection density interrogated), which was the case in seven subjects at 10% and one subject at 15% density. As physiological brain networks should consist of only one connected component (Sotiropoulos & Zalesky, 2019), these sparsest network densities were excluded from our analyses. Therefore, connectivity matrices were thresholded at a range of 20–40% (step size = 1%) network density for each subject.

## 2.6 | GT measures

Node degree and betweenness centrality (BC) were calculated for all nodes using functions available in the Brain Connectivity Toolbox (<https://pypi.org/project/bctpy/>, (Rubinov & Sporns, 2010)). Node



**FIGURE 2** (a) Example data set for one of the two metrics analyzed (degree, betweenness centrality [BC]). A feature-vector of 96-features per subject as well as the respective label (rightward lateralization [RL] dark blue, laterality index [LL] light blue) served as raw input for the classification pipeline. (b) Balancing of classes and (c) train test split prior to (d) feature selection via permutation testing of features across both groups within the training set. (e) Two models were trained and validated on the training set. Via 1000 iterations of fivefold cross-validation using different data-splits in each iteration, confidence intervals and differences to null classifiers for ROC-AUC and accuracy were computed during validation. (f) Testing was conducted on the separate test set, and ROC-AUC and accuracy were computed for model evaluation

degree indicates the sum of connections that link a node to the rest of the network, while BC is a measure of how many shortest paths traverse a node (Bullmore & Sporns, 2009). Such GT metrics summarize topological network properties in a lower dimensionality than, for example, an adjacency matrix does.

GT metrics were computed in each subject and at each network density. This resulted in a curve of the respective metric per node as a function of network density (Figure 1f,g). This curve was integrated to obtain one value of the analyzed metric per node and subject (Bassett et al., 2008), therefore circumventing later statistical testing at multiple density thresholds. For the convenience of the reader, these averaged measures of node degree and BC will remain termed “node degree” and “BC,” respectively.

## 2.7 | Classification of language lateralization

Computation of subject-wise LI resulted in 117 unrelated subjects with RL and 315 subjects with LL. Since the number of samples was small, GT metrics (node degree, BC) were used for model training to avoid overfitting. For each metric, separate data sets were generated,

and thus different models were trained. For an overview of the machine learning procedure, see Figure 2.

In a first step, random undersampling was applied to the majority class (i.e., LL) to numerically match the minority class. Second, 20% of the subjects were randomly split off from the training set to form the test set (Figure 2b,c).

For feature selection, permutation testing with 100k permutations was performed to test for differences in the respective metric across the two groups. Note that this was performed on the down-sampled training set *after* splitting off the test set. All significant features ( $p$  uncorrected  $<.05$ ) were selected in each case for training of the model. Prior to model training, metrics were scaled using a z-transformation. For each metric, a random forest (RF) classifier and a neural network (NN) were trained on the selected features.

The Keras API (<https://keras.io/>) was used to build a simple NN consisting of the input layer, two dense layers with two interspersed dropout layers and a single neuron for output (schematic see Figure 2e). The learning rate was set at 0.001, ReLU was selected as the activation function in each layer and “Adam” was selected as the optimizer. Hyperparameter tuning was conducted using a separate validation set (20% of the initial training set).

Afterward, 1000 iterations of stratified fivefold cross-validation (using 1000 different data splits) were conducted on the entire training set (i.e., 80% of the data) to obtain a mean validation accuracy and AUC-ROC. Confidence intervals were computed, and additional null classifiers were trained with shuffled labels during each iteration, whose performances were compared to the “true” models by computing differences in model performance (AUC-ROC, accuracy) during each iteration. The null hypothesis (classifications do not exceed chance) was rejected if the lower bound of the 95% confidence interval of differences in model performance was greater than zero.

Finally, models were evaluated on the unseen test set. The NN that was trained on node degree, for example, consisted of 29 neurons in the first and 20 neurons in the second dense layer. AUC-ROC and binary accuracy were calculated for model evaluation.

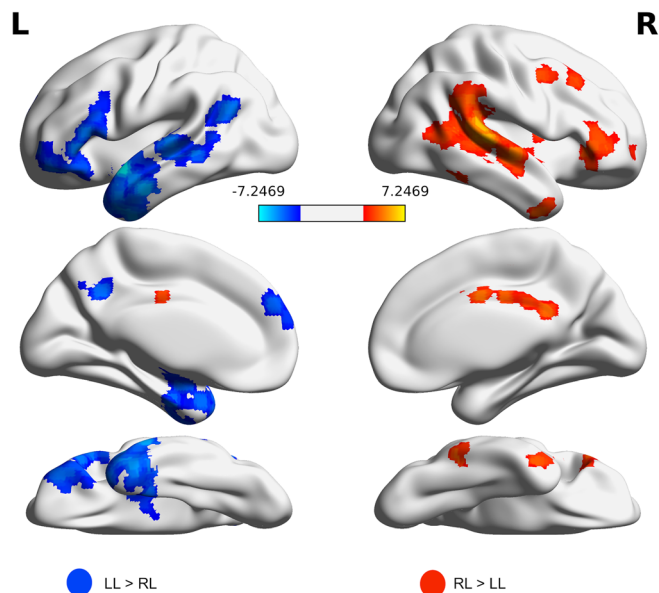
For comparison, we used Scikit-Learn (<https://scikit-learn.org/stable/>) to train an RF classifier on the same data set. Here, the entire training set was used to perform a grid-search for hyperparameter estimation via fivefold cross-validation. Mean validation accuracy and AUC were subsequently obtained via 1000 iterations of stratified fivefold cross-validation as described above, with model hyperparameters set to those of the best estimator during grid search. Finally, the RF was evaluated on the same separate test set as the NN.

Confidence intervals for AUC-ROC were calculated on the test set using a python implementation of the DeLong method (DeLong et al., 1988; Sun & Xu, 2014), [https://github.com/yandexdataschool/roc\\_comparison](https://github.com/yandexdataschool/roc_comparison)). Confidence intervals for model accuracy on the independent test set were obtained via randomly bootstrapping the test set with 1000 bootstrap samples and evaluating model performance for each sample. To assess whether model performance exceeded chance, again, additional “null-models” with identical hyperparameters to those of the “true” models were created. These models were trained on the training set with shuffled data labels and were evaluated on the original test set. The performances of the true and null models were evaluated and compared on the test set by computing differences in model performance for 1000 bootstraps. The null hypothesis (classifications do not exceed chance) was rejected if the lower bound of the 95% confidence interval of differences in model performance was greater than zero.

## 2.8 | Statistical analyses

After model evaluation, statistical analyses were conducted across the entire study population ( $n = 432$  unrelated subjects) for inference of differences in structural connectivity across both groups. The above-mentioned GT measures were compared via permutation testing with 100k permutations and false-discovery rate correction (Benjamini-Hochberg) for multiple testing was applied in each instance.

In addition, link-wise differences in connectivity strength were analyzed with the Network-based Statistics (NBS) Toolbox (Zalesky et al., 2010) using the NBS and FDR correction methods, as recommended. NBS was conducted at densities of 20 and 30% and in



**FIGURE 3** Comparison of task-activations from the Story > Math contrast between subjects with rightward lateralization (RL) and laterality index (LL). Cluster-corrected z-maps of significant activations stronger in subjects with RL are depicted in red, and activations stronger in LL are depicted in blue (here depicted as negative values) at a z-threshold of 3.1. The range of activations for RL > LL was 3.1–6.3 and the scale was adapted to the larger range of z-values from the LL > RL contrast

unthresholded connectomes. Link-wise analyses were followed up by tractometry to explore group-differences in microstructural properties of language-related tracts. These results are reported and discussed in the supplementary material.

## 3 | RESULTS

Stronger RL was found in 125/1038 subjects (12%), and there was no difference in sex across groups (RL: 58.3% female, LL: 53.1% female,  $p$  (uncorrected) = 0.3, OR = 0.81). Then, 97 subjects were left-handed and 26 (26.8%) of which showed RL, while 10.7% of right-handed subjects had been classified as RL. The association between handedness and fMRI-based assessment of language laterality was statistically significant (Fisher's exact test,  $p$ (uncorrected) = .000031, OR = 0.33). Furthermore, there was no difference in task-performance across RL and LL (task accuracy:  $t = 1.355$ ,  $p$ (uncorrected) = .17909,  $d = 0.14$ ; reaction time during task:  $t = 0.31$ ,  $p$ (uncorrected) = .75,  $d = 0.03$ ).

All following analyses were conducted within a subset of 432 unrelated subjects (226 females, no difference in sex across groups, OR = 0.84,  $p = .45$ ), enriched with individuals with RL (117/432 = 27.1%) as detailed in Section 2.1. No difference in task performance across RL and LL was detected in this subset either (task accuracy:  $t = 1.09$ ,  $p$ (uncorrected) = .28,  $d = 0.13$ ; reaction time during task:  $t = -0.17$ ,  $p$ (uncorrected) = .86,  $d = -0.02$ ).

### 3.1 | Differences in task activation

Figure 3 shows that RL was mainly driven by weaker activations in the left ATL compared to subjects with LL, and by stronger activations in the right anterior inferior temporal gyrus and in the right anterior middle temporal gyrus. Concomitant activations stronger in subjects with RL were mainly located in the right posterior middle temporal gyrus (pMTG) and pSTG, while homologous contralateral regions showed significantly decreased activations compared to subjects with LL. Also, right IFG had stronger concomitant activations in RL, with decreased activity in the contralateral IFG compared to LL. The right posterior cingulate cortex (PCC) and right caudal middle frontal gyrus showed stronger concomitant activations in subjects with RL.

### 3.2 | Classification of language lateralization

Classifiers trained on node degree yielded significant predictions of functional lateralization during semantic processing both during cross validation and during testing on the held-out test set. The NN and RF classifier showed similar performances and both significantly outperformed classifiers trained with shuffled data labels. Table 1 depicts

the results obtained from cross validation, as well as from testing of both models on the held-out test set. Predictions of models trained on BC did not exceed chance.

Figure 4 depicts the ROC curves of model performances on the held-out test set, as well as a visualization of features used for classification, while Table 2 lists these features selected for classification with degree-based models. Results obtained using a different method of threshold definition, as well as results obtained using a different cortical parcellation were similar and are reported in the supplementary material.

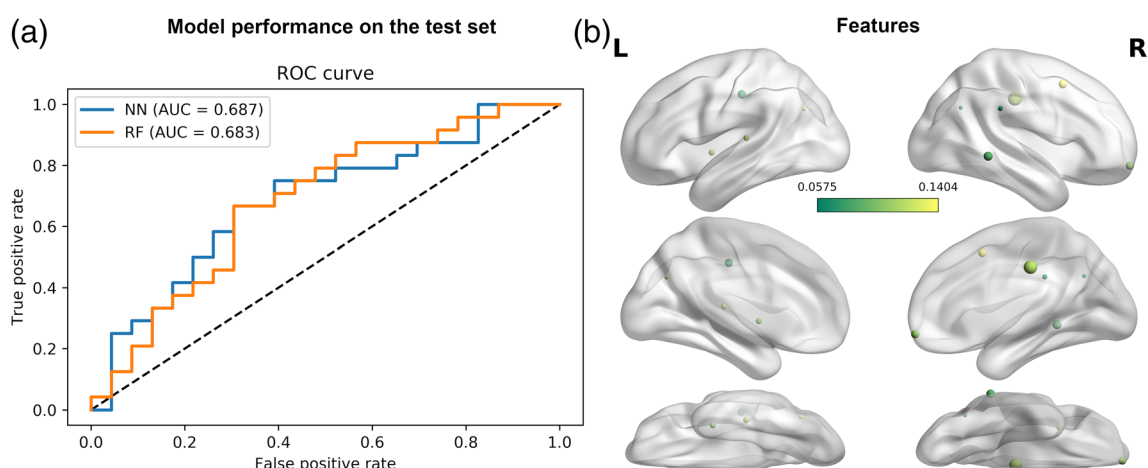
### 3.3 | GT metrics

After model evaluation, GT metrics derived from structural connectomes were analyzed across the entire data set.

Increased node degree was detected within the right pMTG in subjects with RL (pMTG:  $p = .02784$ ,  $d = -0.4$ ,  $t = -3.68$ ), while node degree of the right PCC was greater in subjects with LL ( $p = .04464$ ,  $d = 0.35$ ,  $t = 3.35$ ) (Figure 5). There was no difference in BC in any ROI across groups. All  $p$ -values were FDR-corrected for 96 comparisons corresponding to 96 ROIs. No link-wise differences in connectivity strength were detected.

**TABLE 1** Performance of degree-based classifiers during validation and on the held-out test set. ROC-AUC and accuracies reported for cross-validation are means computed from 1000 models evaluated during 1000 iterations of fivefold stratified cross-validation, each with varying train-test-splits. Both models performed significantly better than chance

	Model	AUC [CI]	Accuracy (%) [CI]	AUC: Difference to null classifiers (M [CI])	Accuracy: Difference to null classifiers (%) (M [CI])
Cross validation	Neural network	0.683 [0.63–0.73]	68.34 [64.2–72.7]	0.14 [0.03–0.26]	10.4 [2–19.1]
	Random forest	0.696 [0.66–0.73]	64.81 [60.9–68.5]	0.19 [0.08–0.3]	14.7 [5.4–24]
Testing	Neural network	0.687 [0.53–0.84]	68.09 [53.2–80.9]	0.36 [0.07–0.62]	24.9 [2.1–46.8]
	Random forest	0.683 [0.527–0.839]	68.09 [55.27–80.85]	0.28 [0.07–0.49]	23.4 [4.3–42.6]

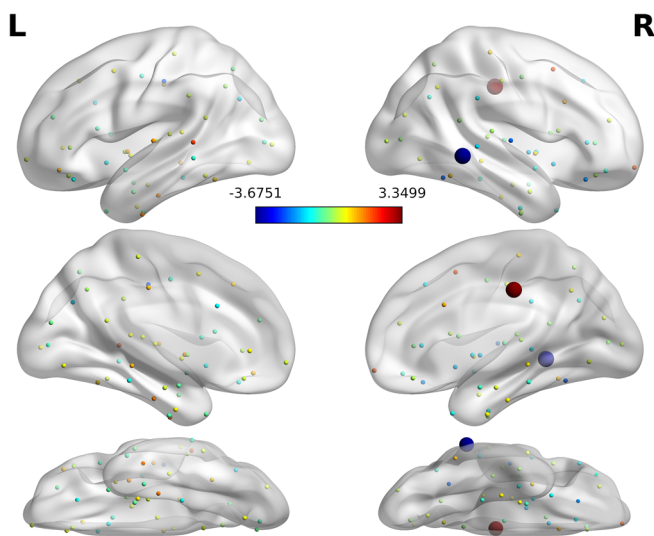


**FIGURE 4** (a) Receiver operating curves obtained during testing of models that had been trained on node degree. The dashed line indicates chance. NN, neural network; RF, random forest. (b) Visualization of features selected for training as well as feature importance as estimated by the random forest and the neural network. Node size indicates feature importance for the neural network, while node color indicates impurity-based feature importance of the random forest



**TABLE 2** Features selected from the degree-based training set for subsequent machine learning (as also depicted in Figure 4). Features were selected via two-tailed permutation testing and negative values indicate greater degree in subjects with RL. *p*-Values are uncorrected

Degree			
Region	<i>t</i> -stat	Cohen's <i>d</i>	<i>p</i>
Right posterior middle temporal gyrus	−3.06	−0.45	.0026
Right caudal middle frontal gyrus	2.69	0.39	.00789
Right posterior cingulate cortex	2.62	0.38	.0094
Right inferior parietal cortex	−2.5	−0.37	.0131
Left postcentral gyrus	−2.32	−0.34	.021
Right frontal pole	2.23	0.32	.027
Right supramarginal gyrus	−2.12	−0.31	.035
Left insula	2.06	0.3	.041
Left posterior transverse temporal gyrus	2.03	0.3	.044
Left inferior parietal cortex	−1.98	−0.28	.049



**FIGURE 5** Results of multiple univariate permutation testing of node degree across groups. Node colors indicate *t*-statistic. Two nodes reached statistical significance (right posterior middle temporal gyrus [pMTG] and right posterior cingulate cortex [PCC]). Two-tailed testing of subjects with laterality index (LL) versus rightward lateralization (RL) was performed, so that negative *t*-statistics indicate stronger degree in subjects with RL

## 4 | DISCUSSION

The utility of structural connectomes for prediction of functional language lateralization in the ATLS was evaluated. In a second step, we provided evidence for a structural basis of this functional lateralization.

### 4.1 | Classification of language lateralization

We demonstrated that better than chance classification of lateralized semantic processing is possible with modest to acceptable AUC–ROC (Mandrekar, 2010) based on dMRI connectomes alone. To our knowledge, this study contributes the first effort to predict hemispheric language dominance from structural connectomes.

The best models were based on node degree, and classifications based on this metric showed robust results with moderate AUC–ROC and accuracy across both classifiers. The confidence intervals of the estimations on the test set were wide, which can be attributed to the small sample size of 234 subjects for training, validation, and testing. The sample size might also have prohibited classifiers from producing better results.

It is noteworthy that some of the predictive regions (e.g., right caudal middle frontal gyrus, right PCC, or right pMTG) also showed stronger concomitant activations in subjects with RL as compared to LL during the language task, as shown in Figure 3, hinting at a structural basis for these rightward lateralized activations in these regions.

Similar results were obtained when using a different, more fine-grained parcellation for generation of connectivity features, among which similar regions were predictive of language laterality, underlining the robustness of our results.

Our findings highlight that dMRI connectivity at rest contains information on task-derived functional language lateralization. Especially the assessment of language laterality in subjects who are unable to perform task-fMRI could benefit from the availability of task-free models, and we believe the present study may pose an encouraging step in this direction.

While previous predictions of language lateralization from resting-state fMRI showed promising results in healthy populations (Tavor et al., 2016), rs-fc seems to be susceptible to effects of, for example, anticonvulsive medication (Salinas & Szabó, 2017; Zhang et al., 2017) or vigilance (Liu & Falahpour, 2020). No such effect is known to affect dMRI, rendering structural connectivity an interesting tool for resting-state assessment of language laterality. dMRI connectome-based predictions involving different or multiple language tasks that elicit perhaps more markedly lateralized activation maps (thus facilitating labeling of data) will be of great interest in the future, especially in clinical populations.

### 4.2 | Graph theoretical analysis

After classification of language lateralization in the ATLS, differences in structural connectivity across the entire study population were analyzed.

Analysis of GT measures showed increased node degree of the right posterior MTG in subjects with RL, which also showed significantly stronger concomitant activations in RL during the language task, while the contralateral posterior MTG showed decreased activations compared to LL. This indicates a concomitant rightward shift of

involvement of the pMTG in semantic processing that is accompanied by measurable increase in structural degree centrality.

Previous studies consistently demonstrated a role of the left pMTG as a major semantic hub where multiple modalities of semantic input converge (Binder et al., 2009; Binder & Desai, 2011; Middlebrooks et al., 2017; Xu et al., 2016). The pMTG was found to be part of a multimodal semantic network consisting of the ventral ATL, PCC, inferior temporal regions, angular gyrus, and mesial temporal regions (Bonilha et al., 2017; Jackson et al., 2016). In patients with left temporal lobe epilepsy, intact fc of the right pMTG was predictive of semantic performance after left ATL resection, indicating a potentially increased role of this contralateral semantic hub in functional reserve (Audrain et al., 2018). Future research should be directed at investigating whether preoperative right pMTG node degree, too, might have prognostic value in determining risk of postoperative language impairment.

Interestingly, the right PCC showed decreased degree in subjects with RL, which showed significantly stronger activations during fMRI in subjects with RL compared to LL. The left PCC was consistently activated in neuroimaging studies of semantic processing and is hypothesized to form an interface to the hippocampus in encoding semantic information to episodic memory (Binder & Desai, 2011). We speculate that our findings might imply a shift of the role of the right PCC toward a more selective function of lower hierarchy (Seguin et al., 2019) in subjects with RL, potentially in conjunction with adopting the above-mentioned function. Interpretation of this finding in relation to semantic processing, however, is not straightforward based on our results and warrants further dedicated research.

Importantly, the node degrees of these two regions (right pMTG and PCC) were among the features selected for classification with the successful degree-based model in the first step of this study.

Using the alternative Schaefer200 parcellation (see supplement), no region withstood correction for 214 comparisons. However, in this exploratory analysis, multiple neighboring parcels along the right posterior MTG showed comparatively marked effect sizes, which would likely have been summed into one ROI (pMTG) in the coarser parcellation, leading to the observed effect. Additional neighboring nodes with increased degree in RL (all nonsignificant after correction for multiple comparisons) were observed within the right posterior STG and in the temporoparietal junction, together forming a cluster over an area resembling a right-sided homologue of the canonical Wernicke's area. The strongest effects were observed in regions within the right caudal middle frontal cortex, which was a predictive region during classification for the models based on the Desikan parcellation as well as for models based on the Schaefer parcellation. Interestingly, there have been prior reports of the potential utility of the (posterior) middle frontal gyrus in functional lateralization of language function in brain tumor patients (Chang et al., 2020; Dong et al., 2016; Gohel et al., 2019). Overall, results from this second parcellation corroborate our findings from the main analysis.

Results remained largely unchanged when applying a slightly different threshold for RL, as well as with the inclusion of all 1038 subjects of the HCP data set with available data.

### 4.3 | Limitations

There are several limitations to this study. First, the labeling of our groups as subjects with "typical" left versus "atypical" right language lateralization had to be determined by the investigators without knowledge of a ground truth. Due to the known, often bilateral activation patterns of the selected language task (Binder et al., 2011), we did not impose more standard (and in this case arbitrary) cutoff values of LI > 0.2 versus < -0.2. Instead, we chose to identify a cutoff value for more rightward laterality from the distribution of LI among all subjects. While selection of this cutoff was still somewhat arbitrary, it resulted in a plausible number of subjects with RL (12%) compared to the literature (~5–7.5% and ~20% in right-handed and left-handed populations, respectively (Knecht, Deppe, et al., 2000; Knecht, Dräger, et al., 2000; Springer et al., 1999; Szaflarski et al., 2002)), and our approach reduced bias and produced meaningful results under permutation testing. Future research should now focus on tasks that produce stronger lateralizations to infer structural differences in subjects with differential language dominance. To our knowledge, large, openly available data sets containing such language tfMRI and dMRI data are not available to date.

Another limitation is the small sample size for classification with machine learning, which created the need to perform feature selection and engineering, leading to a potential loss of information. The reason for this is the scarcity of subjects with RL in a healthy population, and future data sets enriched with individuals with RL could overcome this issue and might produce even better classification results.

Furthermore, the network densities at which connectomes are probed need to be determined by the investigator, which can introduce bias, and results can vary across different density thresholds in the same data set. Thresholding can increase specificity by eliminating spurious connections, but can also lead to loss of information due to elimination of true connections (Sotiropoulos & Zalesky, 2019). To overcome this issue, GT metrics were averaged over a range of connection densities to form a single summary metric per node (Bassett et al., 2008), and thresholded connectomes were rigorously interrogated for fragmentation to ensure anatomical plausibility.

## 5 | CONCLUSION

We report above-chance predictions of functional language lateralization during semantic processing within the ATLS based on structural connectomes. Overall, these results show promise that further efforts with larger sample sizes may achieve even better accuracies in the future.

Structural differences were found in right-sided homologues of widely accepted semantic hubs, one of which (pMTG) adopted a more central role in subjects with RL. Functional implications of the demonstrated decreased topological centrality of the right PCC in rightward lateralized subjects are of great interest and warrant future research. In conclusion, stronger rightward language lateralization in healthy

individuals seems to be supported by a right-sided structural framework involving canonical language regions.

## ACKNOWLEDGMENTS

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University. Open Access funding enabled and organized by Projekt DEAL. WOA Institution: PHILIPPS-UNIVERSITAET MARBURG Consortia Name : Projekt DEAL

## CONFLICT OF INTEREST

Susanne Knake: None related to this work. Susanne Knake received consultant fees or speaker's honoraria from Arvelle, Bial, Epilog, Desitin, Precisis, UCB, and Zogenix. Lars Timmermann: None related to this work. Between January 2018 and January 2021, Lars Timmermann received payments as a consultant for Boston Scientific and also received honoraria as a speaker on symposia sponsored by UCB, Desitin, Boston Scientific, AbbVIE, Novartis, GlaxoSmithKline, and DIA-PLAN. The institution of Lars Timmermann, not Lars Timmermann personally received funding by Boston Scientific, the German Research Foundation, the German Ministry of Education and Research and the Deutsche Parkinson Vereinigung. Neither Lars Timmermann nor any member of his family holds stocks, stock options, patents, or financial interests in any of the above-mentioned companies or their competitors. The other authors declare no competing financial interests.

## DATA AVAILABILITY STATEMENT

The Human Connectome Project data used in this study are freely available under <https://www.humanconnectome.org/>. Data processing was performed with freely available software using FSL (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>), Freesurfer (<https://surfer.nmr.mgh.harvard.edu/>), functions from the Brain Connectivity Toolbox (<https://sites.google.com/site/bctnet/>, <https://github.com/aestrivex/bctpy/>), Keras (<https://keras.io/>), Scikit-Learn (<https://scikit-learn.org/stable/>), and in-house scripts (available at <https://github.com/Zahnert/Semantic-Processing>). Brain images were created using the freely available BrainNet Viewer software (Xia et al., 2013).

## ORCID

Felix Zahnert  <https://orcid.org/0000-0003-0366-6724>

## REFERENCES

- Audrain, S., Barnett, A. J., & McAndrews, M. P. (2018). Language network measures at rest indicate individual differences in naming decline after anterior temporal lobe resection. *Human Brain Mapping, 39*, 4404–4419. <https://doi.org/10.1002/hbm.24281>
- Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., Glasser, M. F., Curtiss, S., Dixit, S., Feldt, C., Nolan, D., Bryant, E., Hartley, T., Footer, O., Bjork, J. M., Poldrack, R., Smith, S., Johansen-Berg, H., Snyder, A. Z., & van Essen, D. C. (2013). Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage, 80*, 169–189. <https://doi.org/10.1016/j.neuroimage.2013.05.033>
- Bassett, D. S., Bullmore, E., Verchinski, B. A., Mattay, V. S., Weinberger, D. R., & Meyer-Lindenberg, A. (2008). Hierarchical organization of human cortical networks in health and schizophrenia. *The Journal of Neuroscience, 28*, 9239–9248. <https://doi.org/10.1523/JNEUROSCI.1929-08.2008>
- Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience, 20*, 353–364. <https://doi.org/10.1038/nn.4502>
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*, 527–536. <https://doi.org/10.1016/j.tics.2011.10.001>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19*, 2767–2796. <https://doi.org/10.1093/cercor/bhp055>
- Binder, J. R., Gross, W. L., Allendorfer, J. B., Bonilha, L., Chapin, J., Edwards, J. C., Grabowski, T. J., Langfitt, J. T., Loring, D. W., Lowe, M. J., Koenig, K., Morgan, P. S., Ojemann, J. G., Rorden, C., Szafarski, J. P., Tivarus, M. E., & Weaver, K. E. (2011). Mapping anterior temporal lobe language areas with fMRI: A multicenter normative study. *NeuroImage, 54*, 1465–1475. <https://doi.org/10.1016/j.neuroimage.2010.09.048>
- Bonilha, L., Hillis, A. E., Hickok, G., den Ouden, D. B., Rorden, C., & Fridriksson, J. (2017). Temporal lobe networks supporting the comprehension of spoken words. *Brain, 140*, 2370–2380. <https://doi.org/10.1093/brain/awx169>
- Bradshaw, A. R., Thompson, P. A., Wilson, A. C., Bishop, D. V. M., & Woodhead, Z. V. J. (2017). Measuring language lateralisation with different language tasks: A systematic review. *PeerJ, 5*, e3929. <https://doi.org/10.7717/peerj.3929>
- Branco, P., Seixas, D., & Castro, S. L. (2020). Mapping language with resting-state functional magnetic resonance imaging: A study on the functional profile of the language network. *Human Brain Mapping, 41*, 545–560. <https://doi.org/10.1002/hbm.24821>
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews. Neuroscience, 10*, 186–198. <https://doi.org/10.1038/nrn2575>
- Bullmore, E. T., & Bassett, D. S. (2011). Brain graphs: Graphical models of the human brain connectome. *Annual Review of Clinical Psychology, 7*, 113–140. <https://doi.org/10.1146/annurev-clinpsy-040510-143934>
- Chang, E. F., Kurteff, G., Andrews, J. P., Briggs, R. G., Conner, A. K., Battiste, J. D., & Sughrie, M. E. (2020). Pure apraxia of speech after resection based in the posterior middle frontal gyrus. *Neurosurgery, 87*, E383–E389. <https://doi.org/10.1093/neuros/nyaa002>
- Cope, T. E., Shtyrov, Y., MacGregor, L. J., Holland, R., Pulvermüller, F., Rowe, J. B., & Patterson, K. (2020). Anterior temporal lobe is necessary for efficient lateralised processing of spoken word identity. *Cortex, 126*, 107–118. <https://doi.org/10.1016/j.cortex.2019.12.025>
- Delgado-Fernández, J., García-Pallero, M. Á., Manzanares-Soler, R., Martín-Plasencia, P., Blasco, G., Frade-Porto, N., Navas-García, M., Pulido, P., Sola, R. G., & Torres, C. V. (2020). Language hemispheric dominance analyzed with magnetic resonance DTI: Correlation with the Wada test. *Journal of Neurosurgery, 134*, 1703–1710. <https://doi.org/10.3171/2020.4.JNS20456>
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics, 44*, 837. <https://doi.org/10.2307/2531595>
- Desai, V. R., Vedantam, A., Lam, S. K., Mirea, L., Foldes, S. T., Curry, D. J., Adelson, P. D., Wilfong, A. A., & Boerwinkle, V. L. (2018). Language lateralization with resting-state and task-based functional MRI in pediatric epilepsy. *Journal of Neurosurgery. Pediatrics, 23*, 171–177. <https://doi.org/10.3171/2018.7.PEDS18162>

- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31, 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- Dong, J. W., Brennan, N. M. P., Izzo, G., Peck, K. K., & Holodny, A. I. (2016). fMRI activation in the middle frontal gyrus as an indicator of hemispheric dominance for language in brain tumor patients: A comparison with Broca's area. *Neuroradiology*, 58, 513–520. <https://doi.org/10.1007/s00234-016-1655-4>
- Doucet, G. E., Pustina, D., Skidmore, C., Sharan, A., Sperling, M. R., & Tracy, J. I. (2015). Resting-state functional connectivity predicts the strength of hemispheric lateralization for language processing in temporal lobe epilepsy and normals. *Human Brain Mapping*, 36, 288–303. <https://doi.org/10.1002/hbm.22628>
- Fernández, G., Specht, K., Weis, S., Tendolcar, I., Reuber, M., Fell, J., Klaver, P., Ruhlmann, J., Reul, J., & Elger, C. E. (2003). Intrasubject reproducibility of presurgical language lateralization and mapping using fMRI. *Neurology*, 60, 969–975. <https://doi.org/10.1212/01.wnl.0000049934.34209.2e>
- Fornito, A., & Bullmore, E. T. (2015). Connectomics: A new paradigm for understanding brain disease. *European Neuropsychopharmacology*, 25, 733–748. <https://doi.org/10.1016/j.euroneuro.2014.02.011>
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J. R., van Essen, D. C., & Jenkinson, M. (2013). The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, 80, 105–124. <https://doi.org/10.1016/j.neuroimage.2013.04.127>
- Gohel, S., Laino, M. E., Rajeev-Kumar, G., Jenabi, M., Peck, K., Hatzoglou, V., Tabar, V., Holodny, A. I., & Vachha, B. (2019). Resting-state functional connectivity of the middle frontal gyrus can predict language lateralization in patients with brain tumors. *AJNR. American Journal of Neuroradiology*, 40, 319–325. <https://doi.org/10.3174/ajnr.A5932>
- Häberling, I. S., Badzakova-Trajkov, G., & Corballis, M. C. (2011). Callosal tracts and patterns of hemispheric dominance: A combined fMRI and DTI study. *NeuroImage*, 54, 779–786. <https://doi.org/10.1016/j.neuroimage.2010.09.072>
- Hernandez-Fernandez, M., Reguly, I., Jbabdi, S., Giles, M., Smith, S., & Sotiropoulos, S. N. (2019). Using GPUs to accelerate computational diffusion MRI: From microstructure estimation to tractography and connectomes. *NeuroImage*, 188, 598–615. <https://doi.org/10.1016/j.neuroimage.2018.12.015>
- Jackson, R. L., Hoffman, P., Pobric, G., & Lambon Ralph, M. A. (2016). The semantic network at work and rest: Differential connectivity of anterior temporal lobe subregions. *The Journal of Neuroscience*, 36, 1490–1501. <https://doi.org/10.1523/JNEUROSCI.2999-15.2016>
- Janecek, J. K., Swanson, S. J., Sabsevitz, D. S., Hammeke, T. A., Raghavan, M., E Rozman, M., & Binder, J. R. (2013). Language lateralization by fMRI and Wada testing in 229 patients with epilepsy: Rates and predictors of discordance. *Epilepsia*, 54, 314–322. <https://doi.org/10.1111/epi.12068>
- Jansen, A., Menke, R., Sommer, J., Förster, A. F., Bruchmann, S., Hempleman, J., Weber, B., & Knecht, S. (2006). The assessment of hemispheric lateralization in functional MRI—Robustness and reproducibility. *NeuroImage*, 33, 204–217. <https://doi.org/10.1016/j.neuroimage.2006.06.019>
- Jbabdi, S., Sotiropoulos, S. N., Savio, A. M., Graña, M., & Behrens, T. E. J. (2012). Model-based analysis of multishell diffusion MR data for tractography: How to get over fitting problems. *Magnetic Resonance in Medicine*, 68, 1846–1855. <https://doi.org/10.1002/mrm.24204>
- Knecht, S., Deppe, M., Dräger, B., Bobe, L., Lohmann, H., Ringelstein, E., & Henningsen, H. (2000). Language lateralization in healthy right-handers. *Brain*, 123(Pt 1), 74–81. <https://doi.org/10.1093/brain/123.1.74>
- Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., Ringelstein, E. B., & Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, 123(Pt 12), 2512–2518. <https://doi.org/10.1093/brain/123.12.2512>
- Li, M., Wu, J., Jiang, P., Yang, S., Guo, R., Yang, Y., Cao, Y., & Wang, S. (2021). Corpus callosum diffusion anisotropy and hemispheric lateralization of language in patients with brain arteriovenous malformations. *Brain Connectivity*, 11, 447–456. <https://doi.org/10.1089/brain.2020.0853>
- Liu, T. T., & Falahpour, M. (2020). Vigilance effects in resting-state fMRI. *Frontiers in Neuroscience*, 14, 321. <https://doi.org/10.3389/fnins.2020.00321>
- Mandrekar, J. N. (2010). Receiver operating characteristic curve in diagnostic test assessment. *Journal of Thoracic Oncology*, 5, 1315–1316. <https://doi.org/10.1097/JTO.0b013e3181ec173d>
- Matsumoto, R., Okada, T., Mikuni, N., Mitsueda-Ono, T., Taki, J., Sawamoto, N., Hanakawa, T., Miki, Y., Hashimoto, N., Fukuyama, H., Takahashi, R., & Ikeda, A. (2008). Hemispheric asymmetry of the arcuate fasciculus: A preliminary diffusion tensor tractography study in patients with unilateral language dominance defined by Wada test. *Journal of Neurology*, 255, 1703–1711. <https://doi.org/10.1007/s00415-008-0005-9>
- Matsuo, K., Chen, S.-H. A., & Tseng, W.-Y. I. (2012). AveLI: A robust lateralization index in functional magnetic resonance imaging using unbiased threshold-free computation. *Journal of Neuroscience Methods*, 205, 119–129. <https://doi.org/10.1016/j.jneumeth.2011.12.020>
- Middlebrooks, E. H., Yagmurlu, K., Szaflarski, J. P., Rahman, M., & Bozkurt, B. (2017). A contemporary framework of language processing in the human brain in the context of preoperative and intraoperative language mapping. *Neuroradiology*, 59, 69–87. <https://doi.org/10.1007/s00234-016-1772-0>
- Murphy, C., Rueschemeyer, S.-A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *NeuroImage*, 147, 19–31. <https://doi.org/10.1016/j.neuroimage.2016.11.067>
- Nowell, M., Vos, S. B., Sidhu, M., Wilcoxon, K., Sargsyan, N., Ourselin, S., & Duncan, J. S. (2016). Meyer's loop asymmetry and language lateralisation in epilepsy. *Journal of Neurology, Neurosurgery, and Psychiatry*, 87, 836–842. <https://doi.org/10.1136/jnnp-2015-311161>
- Parker Jones, O., Voets, N. L., Adcock, J. E., Stacey, R., & Jbabdi, S. (2017). Resting connectivity predicts task activation in pre-surgical populations. *NeuroImage: Clinical*, 13, 378–385. <https://doi.org/10.1016/j.nicl.2016.12.028>
- Piervincenzi, C., Petrilli, A., Marini, A., Caulo, M., Committeri, G., & Sestieri, C. (2016). Multimodal assessment of hemispheric lateralization for language and its relevance for behavior. *NeuroImage*, 142, 351–370. <https://doi.org/10.1016/j.neuroimage.2016.08.018>
- Powell, H. W. R., Parker, G. J. M., Alexander, D. C., Symms, M. R., Boulby, P. A., Wheeler-Kingshott, C. A. M., Barker, G. J., Noppeney, U., Koeppe, M. J., & Duncan, J. S. (2006). Hemispheric asymmetries in language-related pathways: A combined functional MRI and tractography study. *NeuroImage*, 32, 388–399. <https://doi.org/10.1016/j.neuroimage.2006.03.011>
- 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. <https://doi.org/10.1016/j.neuroimage.2012.04.062>
- Rolinski, R., You, X., Gonzalez-Castillo, J., Norato, G., Reynolds, R. C., Inati, S. K., & Theodore, W. H. (2020). Language lateralization from task-based and resting state functional MRI in patients with epilepsy. *Human brain mapping*, 41, 3133–3146. <https://doi.org/10.1002/hbm.25003>
- Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52, 1059–1069. <https://doi.org/10.1016/j.neuroimage.2009.10.003>
- Sabsevitz, D. S., Swanson, S. J., Hammeke, T. A., Spanaki, M. V., Possing, E. T., Morris, G. L., Mueller, W. M., & Binder, J. R. (2003). Use

- of preoperative functional neuroimaging to predict language deficits from epilepsy surgery. *Neurology*, 60, 1788–1792. <https://doi.org/10.1212/01.wnl.0000068022.05644.01>
- Salinas, F. S., & Szabó, C. Á. (2017). Resting-state functional connectivity changes due to acute and short-term valproic acid administration in the baboon model of GGE. *NeuroImage: Clinical*, 16, 132–141. <https://doi.org/10.1016/j.nicl.2017.07.013>
- Sarwar, T., Ramamohanarao, K., & Zalesky, A. (2019). Mapping connectomes with diffusion MRI: Deterministic or probabilistic tractography? *Magnetic Resonance in Medicine*, 81, 1368–1384. <https://doi.org/10.1002/mrm.27471>
- Schaefer, A., Kong, R., Gordon, E. M., Laumann, T. O., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex*, 28, 3095–3114. <https://doi.org/10.1093/cercor/bhx179>
- Seghier, M. L. (2008). Laterality index in functional MRI: Methodological issues. *Magnetic Resonance Imaging*, 26, 594–601. <https://doi.org/10.1016/j.mri.2007.10.010>
- Seguin, C., Razi, A., & Zalesky, A. (2019). Inferring neural signalling directionality from undirected structural connectomes. *Nature Communications*, 10, 4289. <https://doi.org/10.1038/s41467-019-12201-w>
- Silva, G., & Citterio, A. (2017). Hemispheric asymmetries in dorsal language pathway white-matter tracts: A magnetic resonance imaging tractography and functional magnetic resonance imaging study. *The Neuroradiology Journal*, 30, 470–476. <https://doi.org/10.1177/1971400917720829>
- Sotiropoulos, S. N., Jbabdi, S., Xu, J., Andersson, J. L., Moeller, S., Auerbach, E. J., Glasser, M. F., Hernandez, M., Sapiro, G., Jenkinson, M., Feinberg, D. A., Yacoub, E., Lenglet, C., van Essen, D. C., Ugurbil, K., & Behrens, T. E. J. (2013). Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage*, 80, 125–143. <https://doi.org/10.1016/j.neuroimage.2013.05.057>
- Sotiropoulos, S. N., & Zalesky, A. (2019). Building connectomes using diffusion MRI: Why, how and but. *NMR in Biomedicine*, 32, e3752. <https://doi.org/10.1002/nbm.3752>
- Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience*, 17, 652–660. <https://doi.org/10.1038/nn.3690>
- Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S., Brewer, C. C., Perry, H. M., Morris, G. L., & Mueller, W. M. (1999). Language dominance in neurologically normal and epilepsy subjects: A functional MRI study. *Brain*, 122(Pt 11), 2033–2046. <https://doi.org/10.1093/brain/122.11.2033>
- Sun, X., & Xu, W. (2014). Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signal Processing Letters*, 21, 1389–1393. <https://doi.org/10.1109/LSP.2014.2337313>
- Szaflarski, J. P., Binder, J. R., Possing, E. T., McKiernan, K. A., Ward, B. D., & Hammeke, T. A. (2002). Language lateralization in left-handed and ambidextrous people: fMRI data. *Neurology*, 59, 238–244. <https://doi.org/10.1212/WNL.59.2.238>
- Tailby, C., Abbott, D. F., & Jackson, G. D. (2017). The diminishing dominance of the dominant hemisphere: Language fMRI in focal epilepsy. *NeuroImage: Clinical*, 14, 141–150. <https://doi.org/10.1016/j.nicl.2017.01.011>
- Tavor, I., Parker Jones, O., Mars, R. B., Smith, S. M., Behrens, T. E., & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science (New York, N.Y.)*, 352, 216–220. <https://doi.org/10.1126/science.aad8127>
- Tiwari, V. N., Jeong, J., Asano, E., Rothermel, R., Juhasz, C., & Chugani, H. T. (2011). A sensitive diffusion tensor imaging quantification method to detect language laterality in children: Correlation with the Wada test. *Journal of Child Neurology*, 26, 1516–1521. <https://doi.org/10.1177/0883073811409225>
- Tomasi, D., & Volkow, N. D. (2012). Resting functional connectivity of language networks: Characterization and reproducibility. *Molecular Psychiatry*, 17, 841–854. <https://doi.org/10.1038/mp.2011.177>
- Ugurbil, K., Xu, J., Auerbach, E. J., Moeller, S., Vu, A. T., Duarte-Carvajalino, J. M., Lenglet, C., Wu, X., Schmitter, S., van de Moortele, P. F., Strupp, J., Sapiro, G., de Martino, F., Wang, D., Harel, N., Garwood, M., Chen, L., Feinberg, D. A., Smith, S. M., ... Yacoub, E. (2013). Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage*, 80, 80–104. <https://doi.org/10.1016/j.neuroimage.2013.05.012>
- van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., & Ugurbil, K. (2013). The WU-Minn Human Connectome Project: An overview. *NeuroImage*, 80, 62–79. <https://doi.org/10.1016/j.neuroimage.2013.05.041>
- Verhelst, H., Dhollander, T., Gerrits, R., & Vingerhoets, G. (2021). Fibre-specific laterality of white matter in left and right language dominant people. *NeuroImage*, 230, 117812. <https://doi.org/10.1016/j.neuroimage.2021.117812>
- Vernooij, M. W., Smits, M., Wielopolski, P. A., Houston, G. C., Krestin, G. P., & van der Lugt, A. (2007). Fiber density asymmetry of the arcuate fasciculus in relation to functional hemispheric language lateralization in both right- and left-handed healthy subjects: A combined fMRI and DTI study. *NeuroImage*, 35, 1064–1076. <https://doi.org/10.1016/j.neuroimage.2006.12.041>
- Wang, S., van der Haegen, L., Tao, L., & Cai, Q. (2019). Brain functional organization associated with language lateralization. *Cerebral Cortex*, 29, 4312–4320. <https://doi.org/10.1093/cercor/bhy313>
- Wilke, M., & Lidzba, K. (2007). LI-tool: A new toolbox to assess lateralization in functional MR-data. *Journal of Neuroscience Methods*, 163, 128–136. <https://doi.org/10.1016/j.jneumeth.2007.01.026>
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, 14, 1370–1386. <https://doi.org/10.1006/nimg.2001.0931>
- Xia, M., Wang, J., & He, Y. (2013). BrainNet viewer: A network visualization tool for human brain connectomics. *PLoS One*, 8, e68910. <https://doi.org/10.1371/journal.pone.0068910>
- Xu, Y., Lin, Q., Han, Z., He, Y., & Bi, Y. (2016). Intrinsic functional network architecture of human semantic processing: Modules and hubs. *NeuroImage*, 132, 542–555. <https://doi.org/10.1016/j.neuroimage.2016.03.004>
- Zalesky, A., Fornito, A., & Bullmore, E. T. (2010). Network-based statistic: Identifying differences in brain networks. *NeuroImage*, 53, 1197–1207. <https://doi.org/10.1016/j.neuroimage.2010.06.041>
- Zhang, Q., Yang, F., Hu, Z., Zhang, Z., Xu, Q., Dante, M., Wu, H., Li, Z., Li, Q., Li, K., & Lu, G. (2017). Resting-state fMRI revealed different brain activities responding to valproic acid and levetiracetam in benign epilepsy with central-temporal spikes. *European Radiology*, 27, 2137–2145. <https://doi.org/10.1007/s00330-016-4531-z>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Zahnert, F., Kräling, G., Melms, L., Belke, M., Kleinholdermann, U., Timmermann, L., Hirsch, M., Jansen, A., Mross, P., Menzler, K., Habermehl, L., & Knake, S. (2023). Diffusion magnetic resonance imaging connectome features are predictive of functional lateralization of semantic processing in the anterior temporal lobes. *Human Brain Mapping*, 44(2), 496–508. <https://doi.org/10.1002/hbm.26074>