Interhemispheric connections in the maintenance of language performance and prognosis prediction: fully connected layer-based deep learning model analysis

Haosu Zhang, MD, PhD, Kartikay Tehlan, Sebastian Ille, MD, Maximilian Schwendner, MD, Zhenyu Gong, MD, Axel Schroeder, Bernhard Meyer, MD, and Sandro M. Krieg, MD, MBA

Department of Neurosurgery, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Germany

OBJECTIVE Language-related networks have been recognized in functional maintenance, which has also been considered the mechanism of plasticity and reorganization in patients with cerebral malignant tumors. However, the role of interhemispheric connections (ICs) in language restoration remains unclear at the network level. Navigated transcranial magnetic stimulation (nTMS) and diffusion tensor imaging fiber tracking data were used to identify language-eloquent regions and their corresponding subcortical structures, respectively.

METHODS Preoperative image–based IC networks and nTMS mapping data from 30 patients without preoperative and postoperative aphasia as the nonaphasia group, 30 patients with preoperative and postoperative aphasia as the glioma-induced aphasia (GIA) group, and 30 patients without preoperative aphasia but who developed aphasia after the operation as the surgery-related aphasia group were investigated using fully connected layer-based deep learning (FC-DL) analysis to weight ICs.

RESULTS GIA patients had more weighted ICs than the patients in the other groups. Weighted ICs between the left precuneus and right paracentral lobule, and between the left and right cuneus, were significantly different among these three groups. The FC-DL approach for modeling functional and structural connectivity was also tested for its potential to predict postoperative language levels, and both the achieved sensitivity and specificity were greater than 70%. Weighted IC was reorganized more in GIA patients to compensate for language loss.

CONCLUSIONS The authors’ method offers a new perspective to investigate brain structural organization and predict functional prognosis.

https://thejns.org/doi/abs/10.3171/2023.3.FOCUS2363

KEYWORDS interhemispheric connection; deep learning; fully connected layer; nTMS; DTI-FT; language

LANGUAGE-RELATED neural activity can be modulated through reconstruction within widespread networks to maintain function. Information flow across nodes within or between different brain networks via interhemispheric integration and interaction are the central organizational principles for macroscopic brain function. Impairments of interhemispheric connections (ICs) are reportedly related to the decreased functional states in patients with various diseases.

In neurosurgical practice, aphasia may 1) be due to the influence of the tumor preoperatively and 2) arise postoperatively, although these lesions are not in eloquent areas in some cases. Previous studies proposed that the brain connectome concept could be the fundamental organization in maintaining language function and compensating against pathological alterations. According to the current scientific view, the brain attempts to resist the functional burden caused by lesions and assigns cerebral functions to networks in other regions to compensate for function deficits, inevitably involving interhemispheric networks. However, the underlying mechanisms remain unknown.

Previous studies suggested that ICs were essential for...
developing hemispheric linguistic dominance and involved in adaptivity and plasticity. However, it is inconclusive whether these changes were helpful or detrimental to functional restoration. Furthermore, we should also be aware of the unique role of eloquent regions in language networks. Zhang et al. identified regions corresponding to language errors by using navigated transcranial magnetic stimulation (nTMS), which was applied as seeds in diffusion tensor imaging fiber tracking (DTI-FT) to construct brain networks. Therefore, those ICs related to regions identified with nTMS were also investigated in the current study to clarify their roles in language performance.

In addition, predicting postoperative language function remains a significant issue in neurosurgery. Studies by Ille et al. and Zhang et al. proposed that cerebral connectome analysis–based computer techniques can further assist in the interpretation of cerebral structures by improving the accuracy of functional outcomes prediction. Many analyses have already used models, such as convolutional neural networks (CNNs), to analyze medical images for lesion identification and diagnosis. However, selecting the appropriate computational model and techniques for brain network analysis is still a big concern of the medical and computer science communities. In many cases, a complex model is not always the best one. The simple structure of fully connected layers (FCL) in a multilayer perceptron neural network (MLP) allows for ease of use and flexibility. In contrast, convolution layers (CL) in CNN are for image-based analyses, although both FCL and CL belong to deep learning (DL) techniques. Based on FCL, structural locations will be metricized and weighted according to their places in matrices, which can be used as structural contributions in functional analysis. We also compared FCL and CL in the connectomic analysis of the ICs of the present study.

This study analyzed the ICs associated with language function in brain tumor patients by constructing a preoperative interhemispheric DTI-FT that was weighted according to the regions associated with the language errors identified with nTMS. DL methods are optimally suited to perform such data analysis and to help identify the relationship between IC and language function. Our study was based on the hypotheses that 1) nTMS-assisted weighted IC analysis can help distinguish different language function states and 2) a fully connected layer-based DL (FC-DL) model of ICs can be trained and tested for predicting postoperative language function.

**Methods**

**Ethics**

The ethics committee at our hospital approved this study. All procedures were performed in accordance with the Declaration of Helsinki and its subsequent relevant supplements. All subjects were informed about the study and signed the informed consent form.

**Patient Eligibility**

The inclusion criteria were as follows: age older than 18 years; German as their first language; primary left-sided gliomas inside or adjacent to left-sided perisylvian regions or the arcuate fascicles; preoperative MRI available, including T1-weighted images and a diffusion tensor imaging (DTI) scan with 32 diffusion directions; preoperative nTMS language mapping; and no other neurological and/or psychological diseases. Patients with contradictions to nTMS or MRI scanning or other neurological or psychological diseases were excluded. From the data bank of patients who received treatment in our department (2017–2019), 90 patients were enrolled.

Handedness was accessed using the Edinburgh Handedness Inventory, and pathological diagnoses were documented. According to the Aachen Aphasia Test, the aphasia level was rated before the operation and around 1 month postoperatively. In the follow-up analysis, speech status was classified as with or without aphasia. Types of aphasia were not considered. Based on these language states, the patients were grouped as follows: nonaphasia (NA) group (30 cases), with no aphasia before and after the operation; glioma-induced aphasia (GLA) group (30 cases), with preoperative aphasia that was not relieved around 1 month postoperatively; and surgery-related aphasia (SRA) group (30 cases), with normal preoperative language but manifesting aphasia around 1 month postoperatively.

**MRI Data Collection and Language Mapping**

MRI data were collected preoperatively (Achieva 3T, Philips Medical System), including DTI (TR/TE 5000/78 msec, voxel-size 2 × 2 × 2 mm³, 32 diffusion gradient directions, b-value 1000 sec/mm²) together with T1-weighted imaging with contrast (TR/TE 9/4 msec, 1 mm³ isovoxel; 0.5 mmol/ml Dotagraf [Jenapharm GmbH & Co. KG]).

Preoperative language mapping was performed using nTMS (Nexstim xXimia NBS 5.0, Nexstim Plc) following the protocols of previous studies. T1-weighted images were imported and used to navigate to and stimulate the targets that had been predefined by the cortical parcellation system, for which intensity was set to 100% of the resting motor threshold. Regions corresponding to language errors were identified as language-positive regions (POS) by comparing performance on the object-naming task under stimulation versus performance without stimulation.

**Construction of the Interhemispheric Connectome**

First, B0 images were extracted from DTI data as the structural basis. Second, each DTI gradient direction was linearly registered onto the B0 images. Third, the skull was stripped from the T1-weighted images by the HD-bet algorithm (https://github.com/MIC-DKFZ/HD-BET/) and then was linearly registered to the B0 images. Because the POS images were based on T1-weighted imaging files, they were also transformed into B0 space. Then, brain segmentation template AAL90 atlas was deformed for anatomical parcellation of the B0 images.

Next, the constrained spherical deconvolution model from DIPY (Python library DIPY version 1.2.0) was applied for fiber tracking of the ICs. Deterministic FT was performed under a series of fractional anisotropy thresholds that started at 0.0 and increased by 0.01, with the fiber...
length threshold set at 30 mm to exclude U-fibers. Two thresholds for visual ratios (VRs) (≥ 25% and ≥ 50%) were used to filter ICs according to the method proposed by Zhang et al. The connectome consisting of ICs was binarized (connections with > 3 fibers were counted as valid connections and denoted as 1, and the rest were denoted as 0).

Deep Learning

PyTorch version 1.8.0 (https://pytorch.org/; Python 3.8) was applied as a self-supervised DL algorithm that used language levels, tumor locations, and POS. Language-related ICs were extracted from the hidden layers in the FC-DL network on the basis of the whole-brain ICs. An MLP architecture consisting of fully connected linear layers was employed for the DL models. A transfer learning approach was applied, which involved retraining parts of a pretrained network (https://cs231n.github.io/transfer-learning/). The FC-DL network consisted of an auto-encoder and a classifier model. The auto-encoder was used to generate encoding from the first 3 layers that resembled the ICs of the brain (Fig. 1). Then, these ICs were fine-tuned with the transfer learning method to highlight the language-related connections (i.e., weighted ICs) through the classifier model (Fig. 2).

Auto-Encoder Model

An auto-encoder is a neural network trained to replicate its input at its output. The inputs for the auto-encoder were the language-eloquent regions identified with nTMS mapping and the tumor locations as a list of \((N \times 90 \times 1) + (N \times 90 \times 1) = N \times 180 \times 1\) dimensions, where \(N\) was the number of patients in data training, validation, or testing (Fig. 1). The output was the language-eloquent regions as a list of \(N \times 90 \times 1\) dimensions. Encoding comprised an adjacency matrix of connections in the brain with \(N \times 90 \times 90\) dimensions represented as a list of \(N \times 8100\) dimensions. Regularization of the auto-encoder compared the encodings from the third layer to the patient’s existing ICs (Fig. 2) to train layers and to identify intragroup ICs.

Classifier Model

The first 3 layers from the trained auto-encoder model were assigned as the initialization in the classifier model (Fig. 2). The transfer learning–based training process of the classifier model was used to fine-tune the encoding in these layers to emphasize language-related ICs. Regularization of the classifier model calibrated the encoding to the weighted connections within language-eloquent regions, thus finding ICs from language-related connectomes. Then, weighted ICs in the gaussian distribution were extracted. The thresholds for the weights of the ICs were set at 0 and increased to 1 stepwise by 0.01 to calculate summed weights for each group under all thresholds. Specifically, the threshold value of 0.5 for the weighted ICs was identified for intergroup comparison (Fig. 3).

Aphasia was defined as 0 and without aphasia was defined as 1 in the model. During artificial definition, aphasia was defined as 1 and without aphasia as 0 to avoid biases.
FIG. 2. Classifier part of the neural network. The input was the same as in the auto-encoder part (A–C); regularization consisted of weighting the ICs with the nTMS regions to encode the language-related connections in the hidden layers (D–F). The output was the classification of the patient as having aphasia or not postoperatively, which was compared with the available ground truth.

FIG. 3. Language-related ICs. The mean number of summed language-related ICs in each patient group was varied from 0 to 1 as the threshold. The threshold value of 0.5 used in this study (vertical line) is marked to identify the language-related ICs.
Training Process
In the model, the 90 cases were randomly split into training (60% of cases), validation (20%), and test (20%) data sets with K-fold cross-validation (K = 4). As an early stopping criterion, validation losses beyond the limit of patience (epoch number of 50 in this study) was employed to avoid overfitting. Successful training of the FC-DL model was assessed with loss curves.

Statistical Analysis
Ages and tumor sizes were compared using 1-way ANOVA with Tukey correction. The Pearson’s chi-square test was conducted to analyze differences in pathological diagnoses, sex, and handiness. The counts of the weighted ICs based on connectome thresholding at 25% VR and 50% VR were also investigated with 1-way ANOVA to find ICs with intergroup differences. The sensitivity and specificity of the FC-DL method for classification were calculated.

Results

Demographic Analysis
The comparison of demographic information showed no significant differences among the NA, GIA, and SRA groups in terms of sex, handiness, and pathological diagnosis (Table 1). No significant differences in age were detected (NA group 60.7 ± 15.1 years; SRA group 64.2 ± 12.1 years; GIA group 66.9 ± 12.3 years; F = 1.604, p = 0.206).

The tumor size in the NA group was mean ± SD 17.9 ± 18.2 cm³, the GIA group had a tumor size of 34.6 ± 29.8 cm³, and the SRA group had a tumor size of 21.7 ± 22.7 cm³. The intergroup difference was insignificant (p = 0.201).

Connectome Predictions
The weighted adjacency matrix of the ICs for each patient was extracted from the encoding of the first 3 layers of the trained auto-encoder model. These matrices were thresholded at 0.5 to identify the ICs (Fig. 4). Then, after the training process for the classifier model, language-related ICs were identified.

The summed counts of the weighted ICs were significantly greater in the GIA group than in the other patient groups under both the 25% VR (NA group 43.9 ± 26.1; GIA group 57.9 ± 28.0; SRA group 39.9 ± 2.3; F = 11.778, p < 0.001) and 50% VR (NA group 18.2 ± 7.7; GIA group 21.6 ± 7.1; SRA group 16.8 ± 6.6; F = 13.473, p < 0.001) connectomes, with variation in the threshold values ranging from 0 to 1 according to 1-way ANOVA.

ICs between the right paracentral lobule and left pre-

TABLE 1. Comparison of demographic information

<table>
<thead>
<tr>
<th>Item</th>
<th>NA</th>
<th>GIA</th>
<th>SRA</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>6</td>
<td>11</td>
<td>10</td>
<td>0.33</td>
</tr>
<tr>
<td>Female</td>
<td>24</td>
<td>29</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Handedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lt</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0.75</td>
</tr>
<tr>
<td>Rt</td>
<td>26</td>
<td>25</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Pathological diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHO CNS grade 1–3</td>
<td>11</td>
<td>6</td>
<td>10</td>
<td>0.33</td>
</tr>
<tr>
<td>WHO CNS grade 4</td>
<td>19</td>
<td>24</td>
<td>20</td>
<td></td>
</tr>
</tbody>
</table>

* No significant differences in demographic characteristics were noted among the NA, GIA, and SRA patients according to the chi-square test.

FIG. 4. Differences in language-related ICs. Differences in the language-related ICs of three groups of patients are presented as the mean number of the connections in each group. The NA group (A and B), GIA group (C and D), SRA group (E and F) are shown.
cuneus, as well as between the left cuneus and right cuneus, were significantly different among the three groups according to the 1-way ANOVA analysis of the weighted language ICs under both the 25% VR and 50% VR connectomes (Fig. 5, Table 2).

**Classification Model Performance**

The classification tasks had an accuracy of 77.8% on the test data set. The predictions of the classifier model had sensitivity and specificity values greater than 70%. The mean sensitivity was 0.783, and the mean specificity was 0.808 (Table 3).

**Discussion**

In the current study, patients with different language states were compared using the FC-DL model after training with spatial information. Weighted ICs were greatest in the GIA group, followed by the SRA and NA groups. Based on these preoperative data and language prognosis, the FC-DL model can achieve sensibility and specificity.
greater than 70% for predicting postoperative language performance by using weighted ICs.

**nTMS-Based Weighted Interhemispheric Connectome**

According to the nTMS-based weighting method combined with FC-DL, the GIA group had the most ICs, followed by the NA and SRA groups: this suggested a relationship between the tumor-induced aphasic state and enhancement of ICs. SRA patients had proper language function preoperatively but had postoperative aphasia. Based on the present results, it can be inferred that these aphasic states are the stimulator that enhances ICs. In other words, those ICs were an attempt in the brain to compensate for language by strengthening ICs. This suggests that enhancement of ICs may not be solely tumor induced but a remedial measure. But notably, the inclusion of more regions involved in language function could also increase the complexity of the network and consume more time and energy for information processing, potentially leading to a decrease in function. So, these compensatory fibers could be with adverse effects.

Furthermore, the weighted ICs between the right paracentral lobule and left precuneus, as well as between the left and right cuneus, differed significantly between groups. Those regions and their connections are not considered language-eloquent structures according to classic anatomy, in which the distribution of brain functions was considered stable and similar between individuals. Currently, the network tends to be regarded as the functional unit in language. Collaboration and interaction at both the cortical and subcortical levels are considered essential rather than exclusive attributors to cerebral function in several regions. For interpretation, two points should be considered. 1) Critical subcortical structures are involved in language processing as potential functional compensation for the deficits caused by tumors. All patients in this study had tumors in or around language-eloquent regions, and two significant ICs were located in the posterior brain; so, it is possible that these findings were related to patient selection. In other words, ICs in the anterior brain can also be involved in language compensation when tumors are in other locations. 2) It is becoming increasingly accepted that brain networks are the structural basis of language function. These ICs can also be signs of intrahemispheric undercompensation, with the brain seeking interhemispheric involvement.

Moreover, tumor locations and size lacked significant differences among the three groups, but the SRA group had known postoperative aphasia. Taken together, structural reorganization of the SRA group may have focused on the intrahemispheric network rather than the development of support from the contralateral hemisphere as noted in the NA group. In the study by Ille et al., the average shortest path lengths were greater in SRA patients. We speculate that structural reorganization or functional plasticity–related changes occurred in two developing directions: one through intrahemispheric connectome compensation, and the other through more interhemispheric integration. IC changes may appear later than intrahemispheric changes in the ipsilateral hemisphere of the tumor or simultaneously with intrahemispheric changes, but not as primary compensation during the initial phase. More studies are needed to analyze and confirm this.

**DL Network Architecture**

There is still a lack of solid evidence on the specific ICs responsible for language. It was noted that importing all data as input and expecting the language ICs as the output was impossible because there was no available ground truth for the model to enable the learning process. The current study was designed to explore a potential algorithm for further applications combining structural locations and cerebral functions. A variational graph auto-encoder was considered and applied as an unsupervised learning model due to the graphical nature of our data set. The foremost issue was to ensure the correctness of the output from the FC-DL model. Therefore, a solution was to generate the weighted language-related ICs through encoding of the hidden layers of the classifier neural network. The FCL model was applied to embed the cerebral structural information in layers to avoid the problem of structural deviation. Another solution was adding regularization to the auto-encoder loss to reduce overfitting in the classifier model due to the complexity of the neural network architectures. Furthermore, few fully connected linear layers are in a multilayer perceptron; therefore, this was a suitable choice of architecture to enable a reliable classification.

CL are a powerful tool for analyzing image data, such as the LeNet, VGG-Net, and ResNet architectures. However, the attempt to use 2D CL resulted in poor encoding extraction from the hidden layer (Supplementary Fig. 1). The losses also showed instability to converge (Supplementary Fig. 2). Further attempts to feed the input data to the 1D CL failed to find the ICs (Supplementary Fig. 1), and the divergences in training loss implied an inability for classification. Regarding this, the flexibility of FCL in a multilayer perceptron network was suitable for processing 1D data, such as tumor locations and POS, which can be embedded in layers for later classification training.

**MLP-Based Connectome Analysis**

The regularization for encoding the first 3 layers of the auto-encoder model contained the structure information, i.e., the ICs. The weighted nTMS-related connections used as the regularization in the classifier model provided the function information. The metrics of specificity and sensitivity of the classification task showed confidence in the predictions and the language-related ICs. More weighted ICs in the GIA group indicated that the weighted IC outcomes predicted by the FC-DL method were not threshold dependent, further confirming the compensatory attempts of the brain after the loss of language and the effects on structural connectivity due to damage to functions.

Various studies have focused on functional connectivity for brain functions. Martínez et al. demonstrated the role of functional ICs in robustness against damage to cerebral regions, but their study did not include anatomical locations. Pepi et al. analyzed the potential of using machine learning techniques to predict patient epilepsy outcomes on the basis of presurgical EEG connectivity. However, this functional connectivity–focused study also ignored locational information. Carlson et al. also discussed ma-
Recent studies have emphasized the combination of dynamic functional and structural connectivity for understanding the brain processes of language. Deco et al. used dynamic functional and structural connectivity for understanding patients in auditory verbal hallucination studies. The impact of ICs was also identified in this inference. Steinmann et al., in their review, highlighted the consistency of structural and functional interhemispheric connectivity changes in hallucinating patients in auditory verbal hallucination studies. Deligianni et al. also used computational methods to infer functional connectivity from known structural connectivity with diffusion-weighted MRI. In the current study, the transfer learning process in FC-DL provided new insight into the combination of structure and function in ICs.

This study also presented the potential of using DL-based studies to analyze and predict structural and functional modeling. The eloquent linguistic regions in this study were identified with nTMS mapping, which supported the potential application of nTMS for language function tests.

Medical Potential

In this study, several imaging techniques were used to reconstruct structural connections and combined with nTMS to weigh functional regions. Our study was not limited to simple correlations based on certain 1D parameters but instead used complex multidimensional analyses. At present, medical societies have increasingly accepted further integration of artificial intelligence into neurological analysis. The advantage lies in the automatic combination and sourcing of embedded information from data, which can be additionally screened with medical expertise to achieve a reasonable interpretation. This was exactly the analytic approach followed in this study. As we can see, these raw data are present in three dimensions: 1) structural data, including T1-weighted imaging and DTI data; 2) functional data, including nTMS language-related mapping; and 3) brain tumor patients with different language states, which were introduced into the subsequent artificial intelligence analysis. From this, it is evident that medical data analysis will become an essential new interdisciplinary field in the future, requiring at least the integration of medicine, engineering, data analysis, computer science, and other fields.

Limitations

The limitations of the current study should also be noted. First, the study was carried out on a small data set. A larger data set and multicenter study are essentially required for further validation. More demographic parameters should also be included, such as education levels. Second, the current study did not apply postoperative MRI scanning and nTMS. Therefore, postoperative changes in structural connectivity were not included. Third, the postoperative language level was tested only around 1 month after surgery; a longer follow-up should be considered to further identify full language potential.

Conclusions

GIA patients showed more weighted ICs than the others, showing compensatory mechanisms against functional loss. Two ICs beyond classic language regions were greater in the GIA group, indicating that functional compensation was not limited to these regions defined by classic anatomy. Too many connections or regions involved in a particular brain function may increase the cost and complexity during functional implementation and instead cause a decrease in functional performance. The novel application of nTMS-based FC-DL for analyzing functional and structural organization was promising and potentially predicted postoperative language levels.

References


**Disclosures**

Dr. Ille reported personal consulting fees from Brainlab AG and Icotec AG; and honoraria from Carl Zeiss Meditec AG and Nexstim outside the submitted work. Dr. Meyer reported nonfinancial support and personal fees from Brainlab AG and personal fees from Medtronic AG, Icotec AG, Relievant Medsystems, Ulrich, Spineart, and DePuy Synthes. Dr. Krieg reported nonfinancial support and personal fees from Brainlab AG, personal fees from Ulrich and Nexstim, and shares from Need outside the submitted work.

**Author Contributions**

Conception and design: Krieg, Zhang, Tehlan, Schwendner, Meyer. Acquisition of data: Zhang, Ille, Schroeder. Analysis and interpretation of data: Krieg, Zhang, Tehlan, Ille. Drafting the article: Krieg, Zhang, Schwendner. Critically revising the article: Krieg, Zhang, Ille, Schwendner. Reviewed submitted version of manuscript: Krieg, Ille, Schwendner, Meyer. Approved the final version of the manuscript on behalf of all authors: Krieg. Statistical analysis: Zhang, Tehlan. Administrative/technical/material support: Krieg, Zhang, Ille, Schwendner, Meyer. Study supervision: Krieg, Meyer. Figure creation: Gong.

**Supplemental Information**

Online-Only Content
Supplemental material is available online.

[Supplementary Figs. 1 and 2](https://thejn.org/doi/suppl/10.3171/2023.3.Focus2363).

**Correspondence**

Sandro M. Krieg: Technische Universität München, Munich, Germany. sandro.krieg@tum.de.