



Published in final edited form as:

*J Neuroimaging*. 2019 May ; 29(3): 348–356. doi:10.1111/jon.12597.

## Functional MRI Task Comparison for Language Mapping in Neurosurgical Patients

Prashin Unadkat, M.B.B.S.<sup>o</sup>, Luca Fumagalli, M.D.<sup>o</sup>, Laura Rigolo, M.A., Mark G. Vangel, Ph.D., Geoffrey S. Young, M.D., Raymond Huang, M.D., Ph.D., Srinivasan Mukundan Jr., M.D., Ph.D., Alexandra Golby, M.D., and Yanmei Tie, Ph.D.\*

From the Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA (PU, LF, LR, AG, YT); Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA (PU, MG, GSY, RH, SM, AG); School of Medicine and Surgery, Università degli Studi di Milano-Bicocca, Milan, Italy (LF); and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA (MGV).

### Abstract

**Background and Purpose:** Language task-based functional MRI (fMRI) is increasingly used for pre-surgical planning in patients with brain lesions. Different paradigms elicit activations of different components of the language network. The aim of this study is to optimize fMRI clinical usage by comparing the effectiveness of three language tasks for language lateralization and localization in a large group of patients with brain lesions.

**Methods:** We analyzed fMRI data from a sequential retrospective cohort of 51 patients with brain lesions who underwent pre-surgical fMRI language mapping. We compared the effectiveness of three language tasks (Antonym Generation, Sentence Completion, and Auditory Naming) for lateralizing language function and for activating cortex within patient-specific regions-of-interest (ROIs) representing eloquent language areas, and assessed the degree of spatial overlap of the areas of activation elicited by each task.

**Results:** The tasks were similarly effective for lateralizing language within the anterior language areas. The Sentence Completion task produced higher laterality indices within the posterior language areas and had a significantly higher agreement with the clinical report. Dice coefficients between the task pairs were in the range of 0.351 – 0.458, confirming substantial variation in the components of the language network activated by each task.

**Conclusions:** Sentence Completion task consistently produced large activations within the dominant hemisphere and was more effective for lateralizing language within the posterior language areas. The low degree of spatial overlap among the tasks strongly supports the practice of using a battery of tasks to help the surgeon avoid eloquent language areas.

\*Corresponding Author: Yanmei Tie, Ph.D., Assistant Professor, Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Hale Building for Transformative Medicine, 8016G, 60 Fenwood Road, Boston, MA 02115, USA, ytie@bwh.harvard.edu, Tel: 617-732-8249.

These Authors contributed equally to this work

## Keywords

BOLD fMRI; language mapping; pre-surgical planning; brain lesions; brain tumors

---

## Introduction

Preservation of neurologic function correlates strongly with quality of life and overall survival,<sup>1</sup> so it is a critical goal of neurosurgical planning. Blood oxygen level-dependent functional magnetic resonance imaging (BOLD fMRI) allows non-invasive identification of eloquent cortex involved in language, motor, visual and sensory functions, and it is increasingly used in pre-surgical risk assessment and planning to reduce the rate of postoperative functional deficits.<sup>2</sup>

Language function involves a complex network of brain structures with widespread representation in both hemispheres.<sup>3</sup> The use of multiple paradigms has been advocated for clinical purposes<sup>4</sup> because the location and degree of recruitment of different cortical structures vary with both individual functional anatomy and fMRI paradigm.<sup>5,6</sup> Understanding the differences between paradigms is essential for rationally selecting a standardized battery of tasks. In 2017, in the light of evidence available on the commonly used paradigms,<sup>7-9</sup> the American Society of Functional Neuroradiology published two sets of language tasks that balanced ease of application and clinical usefulness.<sup>10</sup> The recommendation for adult clinical fMRI comprises Sentence Completion, Silent Word Generation, Rhyming, Object Naming, and/or Passive Story Listening. The pediatric recommendation consists of Sentence Completion, Rhyming, Antonym Generation, or Passive Story Listening. Despite this, striking variability persists across centers.<sup>11,12</sup>

For brain lesions close to the putative language areas, the risk of post-operative deficits significantly increases when surgery involves the dominant hemisphere, thus it is important to determine the dominant hemisphere for surgical planning.<sup>13,14</sup> Previous studies comparing different tasks have focused on the determination of the language lateralization in healthy subjects<sup>15</sup> and epilepsy patients.<sup>16,17</sup> One study in brain tumor patients suggested that the expressive paradigms gave more valid results for language lateralization.<sup>8</sup> Other studies comparing tasks for language localization used only healthy subjects' data.<sup>7,18,19</sup> The only study that compared three tasks in brain tumor patients focused on the inter-observer variability for the interpretation of the fMRI language maps.<sup>20</sup>

In this study, we evaluated the variation among three widely used tasks in their effectiveness for localizing and lateralizing language function in 51 neurosurgical patients with focal brain lesions. Spatial overlap between the language maps was also assessed. Our goal was to identify the similarities and differences across the language maps derived from these tasks in individual patients, in order to better interpret pre-surgical fMRI results. In fact, a more thorough understanding of the activation patterns of these paradigms can help decide which task to prioritize, or which one should be considered more robust. To our knowledge, this is the first comparative study of fMRI language task paradigms in a large cohort of patients with brain lesions.

## Methods

### Patients

We retrospectively examined all fMRI data from the patients who underwent language fMRI at our institution (Brigham and Women's Hospital, Boston, USA) between September 2014 and July 2016 (228 patients); 59 patients performed all three language tasks, i.e., Sentence Completion (SC), Antonym Generation (AG), and Auditory Naming (AN). Eight patients had non-diagnostic scans due to motion or other artifacts (6 patients) or poor task compliance (2 patients). The remaining 51 patients were included in this study (29 males, 22 females, mean age =  $49.33 \pm 16.3$  years (mean  $\pm$  standard deviation), range: 21–78 years).

The population included patients with primary brain tumors (27 patients including 1 meningioma, 6 low-grade gliomas, and 20 high-grade gliomas), metastases (4 patients), vascular malformations (4 patients), infections (1 patient), and cortical dysplasia or malformations (15 patients). Thirty-eight lesions were in the left hemisphere and 13 lesions were in the right hemisphere. Among the 31 patients with either primary or metastatic brain tumors, 24 patients had left-sided tumors (8 frontal, 9 temporal, 4 parietal, 1 occipital, and 2 insular), and 7 patients had right-sided tumors (3 frontal, 1 temporal, 2 parietal, and 1 intraventricular). All patients were English speakers. Forty patients were right-handed, 8 were left-handed, and 3 were ambidextrous. Handedness was determined by the Edinburgh Handedness Inventory (EHI)<sup>21</sup> in 42 patients. In the 9 patients who did not complete the EHI, handedness was retrieved from clinical records. The Partners Institutional Review Board approved the study, and all patients provided written informed consent.

### Image acquisition

fMRI was performed on 3.0 Tesla Scanners (Siemens Verio, Trio & Skyra Systems, Munich, Germany). Acquisition parameters were optimized for each individual scanner hardware and software. Single-shot gradient echo-planar imaging (EPI) was used to acquire BOLD functional images (repetition time (TR) = 2000 ms, echo time (TE) = 30–40 ms, flip angle = 85–90°, no slice gap, FOV = 22–25.6 cm, acquisition matrix =  $64 \times 64$  or  $128 \times 128$ , voxel size =  $1.88 \times 1.87 \times 4$  or  $1.88 \times 1.87 \times 5$  mm<sup>3</sup>, 24–33 axial slices, ascending interleaved sequence). During the same scanning session, T2-weighted 3D axial images were routinely acquired for better visualization of non-enhancing lesions (repetition time (TR) = 2000–3000 ms, echo time (TE) = 232–448 ms, flip angle = 120°, matrix =  $256 \times 256$  or  $320 \times 320$ , voxel size =  $0.8 \times 0.8 \times 1.2$  or  $1 \times 1 \times 1$  mm<sup>3</sup>, 144–192 axial slices).

### Behavioral paradigms

All patients performed three block-design language tasks (Table 1) and were asked to give non-vocalized responses to limit head motion. The SC task is a visually presented paradigm which stimulates word retrieval, reading comprehension, attentional and semantic processes, and recruits both receptive and expressive language function.<sup>22</sup> The AG task is a visually presented paradigm which stimulates phonologic, working memory, lexical search, semantic and orthographic processes involved with speech and language function.<sup>10,23</sup> The AN task is an auditorily presented paradigm which recruits receptive and expressive language regions,

in addition to task-specific auditory cortex within Heschl's gyrus.<sup>24</sup> In our clinical practice, these tasks were most often presented in a pre-determined order: SC, AG, and AN.

Task paradigms were displayed using an FDA-approved hardware (NordicNeuroLab, Bergen, Norway), and the stimuli were synchronized with the MRI trigger pulse.

### Image processing

MR images were processed using NordicBrainEx software (NordicNeuroLab, Bergen, Norway). Functional images were co-registered to the 3D T2-weighted images and visually inspected for accuracy. This was followed by motion correction, spatial smoothing (6 mm full-width-half-maximum (FWHM) Gaussian kernel), and high-pass filtering (128-second). The general linear model (GLM) approach was applied to analyze the data.<sup>25</sup> In the first-level single subject analysis, an estimate of the canonical hemodynamic response function (HRF) was used as the basis function with the task and baseline conditions explicitly modeled. Following current guidelines,<sup>26</sup> functional maps were thresholded for clinical interpretation by one neuroscientist for all patients.

### Generation of patient-specific language regions-of-interest (ROIs)

Patient-specific language ROI were generated from brain regions corresponding to the canonical anatomical regions of Broca's and Wernicke's areas.<sup>3,27,28</sup> The anterior language ROI was defined as the union of pars opercularis, pars triangularis, and pars orbitalis of the left inferior frontal gyrus (Brodmann Areas (BAs) 44, 45 and 47). The posterior language ROI was defined as the union of angular gyrus, supramarginal gyrus, and the posterior half of the superior temporal gyrus of the left temporoparietal cortex (BAs 22, 39 and 40). Homologous ROIs in the right hemisphere were also created.<sup>29</sup> The ROIs were defined in the Montreal Neurological Institute (MNI) space and constructed from the Talairach Daemon database,<sup>30</sup> using the WFU Pick Atlas software (Version 1.04, Wake Forest University, Winston-Salem, NC, USA).<sup>31</sup> A "reverse" spatial normalization procedure<sup>32</sup> was then performed on these ROIs to generate the patient-specific language ROIs, using the SPM12 software (Wellcome Trust Center for Neuroimaging, University College London, London, UK) (Figure 1 shows an example patient's ROIs). Briefly, the T2 image of each patient was fed into the Segment module in SPM12, where an inverse matrix was derived from the spatial normalization procedure embedded in the module. This inverse matrix was used to transform the standard ROI images into the patient's own brain space. The resulting patient-specific ROIs were then overlaid onto the patients' structural image to confirm their structural accuracy by visual inspection.

### Comparison metrics

The activation volume within each language ROI was measured for each task using 3D Slicer software<sup>33</sup> ([www.slicer.org](http://www.slicer.org)), and then normalized to the total volume of the corresponding ROI to generate a normalized activation volume (NAV) for comparison across patients and tasks.

The laterality index (LI), which represents relative recruitment of the left and right hemispheres towards language function,<sup>17</sup> was calculated separately within the anterior and posterior language ROIs (Equation 1).

$$LI = \frac{NAV_L - NAV_R}{NAV_L + NAV_R} \quad (\text{Equation 1})$$

where  $NAV_L$  refers to the NAV within the left ROI and  $NAV_R$  refers to the NAV within the right ROI. Based on the LI results, language lateralization was classified as left ( $LI \geq 0.2$ ), right ( $LI \leq -0.2$ ), or bilateral ( $-0.2 < LI < 0.2$ ).<sup>34</sup> The agreement between the language lateralization from the fMRI maps and that from the clinical report was assessed. The clinical report was generated by a neuroradiologist who inspected all three language maps separately and reported an overall clinical interpretation of language lateralization for each patient.

To evaluate the spatial overlap between the language maps, the Dice coefficient was calculated within the anterior and posterior ROIs separately for each task pair (Equation 2).<sup>35</sup>

$$\text{Dice coefficient} = \frac{2 * (V_{\text{overlap}})}{V_1 + V_2} \quad (\text{Equation 2})$$

where  $V_{\text{overlap}}$  represents the overlapping activation volume of the task pair, and  $V_1$  and  $V_2$  refer to the activation volume of task 1 and 2, respectively. The Dice coefficients were calculated only for the dominant hemisphere in those patients for whom all three language maps agreed on language laterality (43 patients in the anterior ROI, and 26 patients in the posterior ROI).

## Statistics

For each measurement (NAV, LI, and Dice coefficient), the non-parametric Friedman test<sup>36</sup> was performed to compare across the three tasks using R<sup>37</sup> and Stata software (StataCorp LLC, College Station, TX, USA). We also performed post-hoc pairwise comparison test for each task pair using Friedman's pairwise test when the Friedman test was significant across all tasks. We further performed McNemar's test to assess the agreement between the LI results from the fMRI maps and the clinical report.<sup>38</sup> We considered  $p < 0.05$  as statistically significant, and all between-task pairwise comparisons were corrected for multiple comparisons of the three task pairs using Bonferroni correction.

## Results

The patients' task performance was evaluated by a neuroscientist during pre-scan training, and during scanning by assessing their alertness and responsiveness in between fMRI runs. Thirty-nine patients had no issues with task performance; 7 had mild difficulty, defined as a few missed responses; and 5 had moderate difficulty, defined as some difficulty with active

components of the task (such as difficulty in producing responses without issues with reading or listening to the stimuli). All patients' head motion was at acceptable levels with maximum displacement less than 2 mm in the x, y and z axes.

### Language localization

Figure 2 (A to D) plots all patients' individual tasks' NAVs within the four language ROIs, with the patients ordered according to their NAVs averaged across the three tasks. The NAVs were significantly different among the three tasks in all four ROIs ( $p < 0.001$  for the left anterior, left posterior, and right posterior ROIs,  $p = 0.02$  for the right anterior ROI). Figure 2 (E to H) shows the box plots of the three tasks' NAVs within the four language ROIs, and Table 2 shows the Bonferroni-corrected p value of each task pair comparison and the effect size. Within the left anterior ROI (Figure 2E), both the AG and SC tasks had significant larger NAVs than the AN task, without significant difference between the AG and SC tasks. Within the left posterior ROI (Figure 2F), both the SC and AN tasks had significant larger NAVs than the AG task, without significant difference between the SC and AN tasks. Within the right anterior ROI (Figure 2G), there was no significant difference in the NAVs between the task pairs after Bonferroni correction. Within the right posterior ROI (Figure 2H), the AN task produced significant larger NAV than the AG task, without significant difference between the AG and SC tasks.

Figure 3 shows the box plots of the three tasks' NAVs within the language ROIs in the dominant hemisphere, and Table 3 shows the Bonferroni-corrected p value of each task pair comparison and the effect size. In the dominant anterior ROI (Figure 3A), both the AG and SC tasks had significant larger NAVs than the AN task, without significant difference between the AG and SC tasks. In the dominant posterior ROI (Figure 3B), both the SC and AN tasks had significant larger NAVs than the AG task, without significant difference between the SC and AN tasks.

Dice coefficients between the task pairs in the dominant hemisphere ranged between 0.351 – 0.458 (Table 4), and were not statistically different across all task pairs ( $p > 0.05$ ). Figure 4 shows an example patient's language fMRI maps overlaid on the structural image.

### Language lateralization

Figure 5 (A and B) plots all patients' individual tasks' LIs within the anterior and posterior language ROIs, with the patients ordered according to their LIs averaged across the three tasks. Table 5 shows the Bonferroni-corrected p value of each task pair comparison and the effect size. Within the anterior language ROI (Figure 5A), there was no statistical difference among the LIs of the language maps generated from the three tasks ( $p = 0.44$ ). Within the posterior language ROI (Figure 5D), the SC task had the highest LI compared with the AG task and the AN task without significant difference between the AG and AN tasks.

Within the anterior ROI, McNemar's test demonstrated no significant difference between the LIs of the language maps and the clinical report that was considered as the pseudo ground truth for language lateralization (SC/AG/AN = 90.2%/84.3%/86.3%,  $p > 0.05$  for all task pairs). Within the posterior language ROI, the SC task had a statistically higher agreement rate with the clinical report (80.4%) than the other two tasks (AG = 60.8%,  $p < 0.01$  and AN

= 51%,  $p < 0.01$ ), with the AN task having the lowest agreement rate with the other two tasks (Table 6).

## Discussion

In this retrospective study, we compared the effectiveness of three fMRI tasks for language localization and lateralization in a large number of neurosurgical patients in order to optimize paradigm selection and fMRI interpretation for surgical decision-making in patients with focal brain lesions. We included a large sequential population of patients with brain tumor and other neurosurgical conditions.

Selection of the optimal language task paradigms remains a matter of debate in clinical fMRI,<sup>7–10</sup> because each paradigm yields slightly different results due to differences in sensory input modality, task complexity, and specific aspects of language processing elicited.<sup>5,6</sup> Studies have compared and quantified these differences for various language tasks in healthy volunteers. Zacà et al.<sup>7</sup> and Binder et al.<sup>39</sup> compared a battery of language tasks in healthy volunteers to evaluate their performance in localization and lateralization of language function. Informed by this research, we chose a battery comprising two visual presentation tasks (SC and AG) and one auditory presentation task (AN). This is essential in patients who may have different sensory deficits. In addition, the AG task relies on word-level language processing, and the SC task relies on sentence-level language processing; therefore, we could study both semantic and syntactic processes of language function.

Interpretation of clinical fMRI for pre-surgical planning is further complicated by patient-specific alterations in the functional activation patterns resulting from the lesion itself, including neurovascular uncoupling in high-grade glioma patients, and neuroplasticity in patients with chronic or slow growing lesions with the recruitment of the non-lesional hemisphere.<sup>7,40,41</sup>

### Language localization

A study of 10 male healthy subjects demonstrated that results vary between the tasks imposing greater or lesser cognitive demands.<sup>6</sup> This is likely to be even more significant in lesional patient populations with compromised cognitive function. In the anterior ROI, the AN task produced a smaller NAV, compared to the AG and SC tasks. This raises the possibility that processing transient auditory stimuli imposes greater cognitive demands, resulting in suboptimal stimulation of the language network in lesional patients. Fatigue related to the fact that AN was generally presented last may have exacerbated this effect in our patient cohort.

The finding that the AG task produced a smaller NAV compared with the SC and AN tasks in the posterior ROI may reflect the fact that the word-level AG task may not entail as much syntactical processing as the sentence-level SC and AN tasks.<sup>23</sup> In addition, it is likely that the AN-related primary auditory cortex activation contributed to the larger NAV in the bilateral posterior language ROIs.

## Language lateralization

It is worthy to note that the LI calculated based on the number of activated voxels can be misleading in patients with very low level of activation because weak activation in one hemisphere can produce a high LI when there is no activation within the contralateral hemisphere. In such cases, the high LI may reflect the use of a poorly-suited task for mapping function in a given region, or reflect suboptimal task performance.

Within the anterior language ROI all tasks produced statistically similar LIs. Within the posterior ROI, the SC task produced higher LIs, while the AN task produced the most bilateral language maps, possibly due to the inclusion of the adjacent bilateral primary auditory cortex activation. This is consistent with the reports that auditory tasks inherently induce extensive bilateral temporal cortex activation and more overall activation of the whole cortex, while reading elicits more lateralized activation.<sup>5</sup>

The SC task had the highest agreement with the clinical report within the anterior and posterior language ROIs. One possible explanation is that the neuroradiologists may have placed greater emphasis on the SC task compared to the AG and AN tasks when reporting the clinical interpretation of language lateralization. This is plausible because the SC task almost always agreed with at least one of the other two tasks and consistently produced strongly lateralized results. Such concordance would carry greater weight in clinical interpretation when the reader is required to synthesize information from different tasks and arrive to an overall interpretation of language lateralization.

## Spatial overlap across tasks

The task pairs showed only fair spatial overlap, 4 which is consistent with the literature documenting substantial differences in the language network components activated by different tasks.<sup>6,42</sup> This result supports the rationale for using a comprehensive battery of different tasks to increase the sensitivity of fMRI language mapping, especially in patients with focal lesions and/or pre-operative functional deficits. For example, patients with extreme refractory errors or visual field cuts can generally perform the AG (single word) or AN (auditory stimuli) tasks, but often have difficulty in performing the SC task, which requires greater visual perception and larger visual field. Additionally, in the event of technical failures of one task, the availability of other tasks is of great importance to yield interpretable results for pre-surgical language mapping.

## Limitations

There are some inherent limitations of this study. First, only one neuroscientist manually selected the t-score for thresholding the fMRI maps, which may introduce subjectivity. In our practice, the t-threshold was selected by reviewing the functional maps over multiple voxel-wise statistical significance thresholds and selecting a map for optimal activation visualization based on prior anatomic knowledge. This procedure is in agreement with the current guidelines by the American College of Radiology.<sup>26</sup> Additionally, selecting higher t-scores tends to give a more lateralized functional map.<sup>43</sup> To mitigate the effect of thresholding, recent studies have proposed threshold-independent approaches for LI calculation, however not widely used in clinical practice.<sup>9,44</sup> As our study focused on the



clinical use of language fMRI, we used the t-score that was selected by the same neuroscientist and reported for clinical implementation. Second, while the patient population was selected to allow us to investigate real-world application of fMRI in lesional patients, some lesions involving the language pathways may have altered the patients' ability to process language and, therefore, perform tasks. All patients underwent a pre-scan training period in order to reduce bias related to impaired task performance allowing a straightforward statistical analysis. We chose to include only those patients who could follow and respond to all three tasks during the pre-scan training period, because it would be difficult to draw conclusions on the utility and robustness of the various tasks in the patients whose fMRI was of limited or non-diagnostic quality. Third, the order of the task presentation was not counterbalanced, which may result in suboptimal performance of the AG and AN tasks due to progressive patient fatigue. Fourth, our patient population was heterogeneous with regards to pathology and lesion location. This may have influenced our results, since the lesions may have differently altered the fMRI signals, depending on the nature of the pathology and its proximity to the language areas. However, this is an inherent limitation of clinical fMRI studies. In our study, we have accumulated a large database of the patients who underwent the same series of language fMRI tasks, and we expect that this represents the real-world scenario of clinical application of fMRI.

## Conclusions

By comparing three widely used language tasks in a large cohort of lesional neurosurgical patients, we found that the SC task produced the largest activation volume within the dominant hemisphere, all tasks generated similar LI within the anterior language ROI, and the SC task produced higher LI within the posterior language ROI. The low Dice coefficients confirmed that there was substantial variability in the language network components elicited by different tasks, reinforcing the necessity of including several tasks in pre-surgical fMRI protocols to increase the sensitivity of fMRI language mapping.

## Acknowledgements and Disclosure:

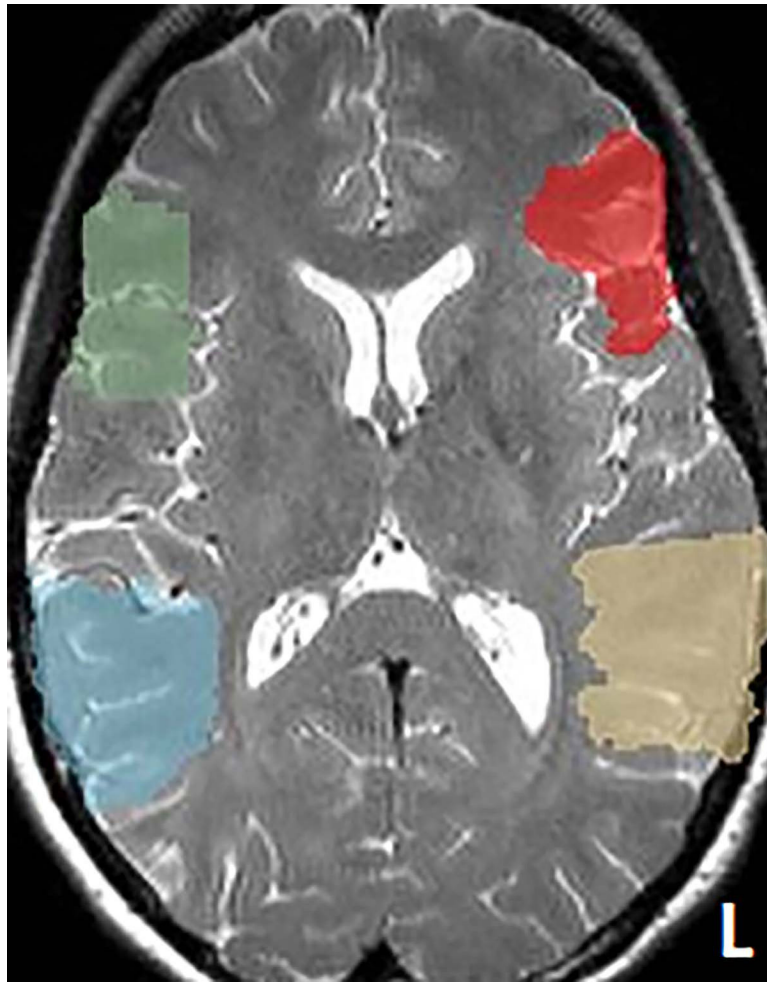
This work is supported by the funding from the National Institutes of Health (NIH) through Grants R21NS075728, R21CA198740, P41EB015898, P41RR019703, and R25CA089017. We confirm that this work is original and has not been published or under consideration for publication elsewhere. We have no conflicts of interest to disclose.

## References

1. McGirt MJ, Mukherjee D, Chaichana KL, et al. Association of surgically acquired motor and language deficits on overall survival after resection of glioblastoma multiforme. *Neurosurgery* 2009;65:463–9. [PubMed: 19687690]
2. Kundu B, Penwarden A, Wood JM, et al. Association of functional magnetic resonance imaging indices with postoperative language outcomes in patients with primary brain tumors. *Neurosurg Focus* 2013;34:E6.
3. Binder JR, Frost JA, Hammeke TA, et al. Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci* 1997;17:353–62. [PubMed: 8987760]
4. Benjamin CF, Walshaw PD, Hale K, et al. Presurgical language fMRI: Mapping of six critical regions. *Hum Brain Mapp* 2017;38:4239–55. [PubMed: 28544168]
5. Buchweitz A, Mason RA, Tomitch LMB, Just MA. Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension. *Psychol Neurosci* 2009;2:111–23. [PubMed: 21526132]

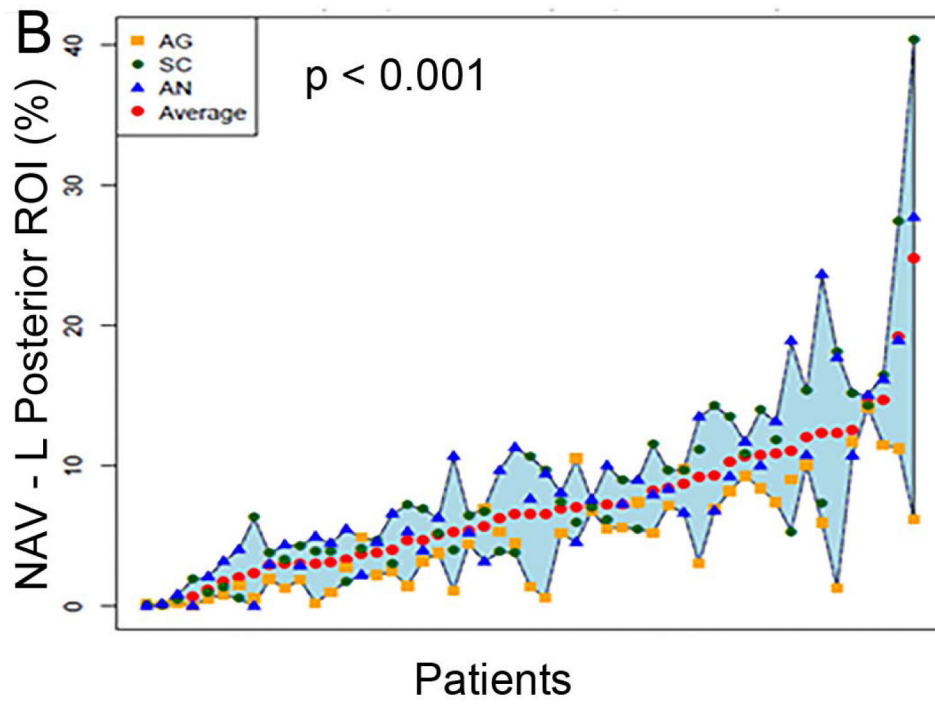
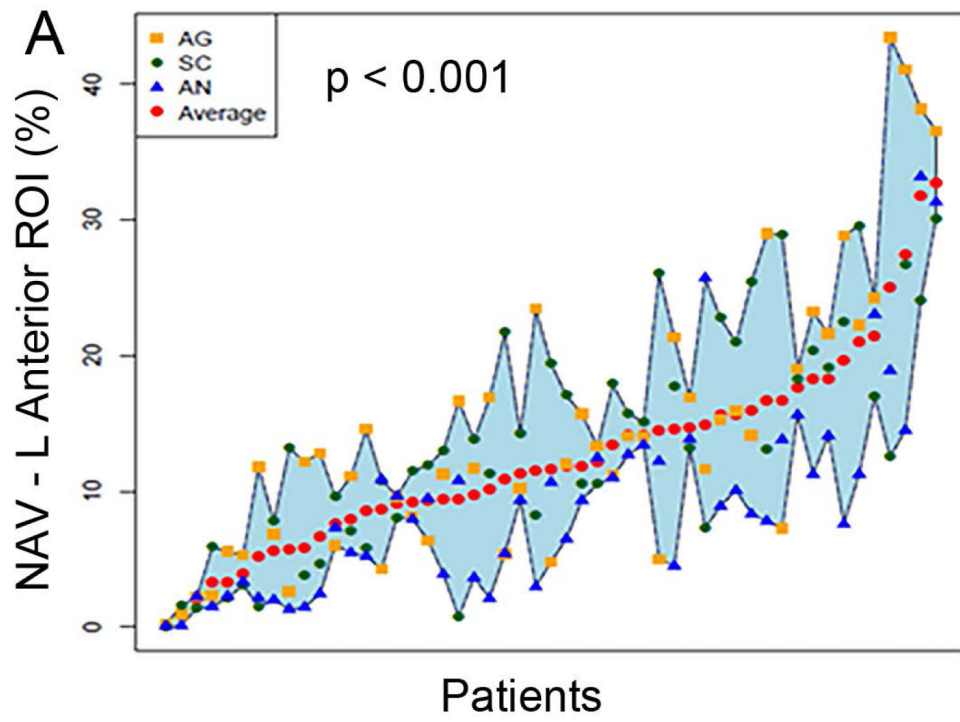
6. Jobard G, Vigneau M, Mazoyer B, Tzourio-Mazoyer N. Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage* 2007;34:784–800. [PubMed: 17110132]
7. Zacà D, Jarso S, Pillai JJ. Role of semantic paradigms for optimization of language mapping in clinical fMRI studies. *Am J Neuroradiol* 2013;34:1966–71. [PubMed: 23788599]
8. Zacà D, Nickerson JP, Deib G, Pillai JJ. Effectiveness of four different clinical fMRI paradigms for preoperative regional determination of language lateralization in patients with brain tumors. *Neuroradiology* 2012;54:1015–25. [PubMed: 22744798]
9. Pillai JJ, Zacà D. Relative utility for hemispheric lateralization of different clinical fMRI activation tasks within a comprehensive language paradigm battery in brain tumor patients as assessed by both threshold-dependent and threshold-independent analysis methods. *Neuroimage* 2011;54:S136–45. [PubMed: 20380883]
10. Black DF, Vachha B, Mian A, et al. American society of functional neuroradiology-recommended fMRI paradigm algorithms for presurgical language assessment. *Am J Neuroradiol* 2017;38:E65–73. [PubMed: 28860215]
11. Bookheimer S Pre-surgical language mapping with functional magnetic resonance imaging. *Neuropsychol Rev* 2007;17:145–55. [PubMed: 17484055]
12. Pillai JJ. The evolution of clinical functional imaging during the past 2 decades and its current impact on neurosurgical planning. *Am J Neuroradiol* 2010;31:219–25. [PubMed: 20150316]
13. Knecht S Handedness and hemispheric language dominance in healthy humans. *Brain* 2000;123:2512–8. [PubMed: 11099452]
14. Bonelli SB, Thompson PJ, Yogarajah M, et al. Imaging language networks before and after anterior temporal lobe resection: Results of a longitudinal fMRI study. *Epilepsia* 2012;53:639–50. [PubMed: 22429073]
15. Ramsey NF, Sommer IE, Rutten GJ, Kahn RS. Combined analysis of language tasks in fMRI improves assessment of hemispheric dominance for language functions in individual subjects. *Neuroimage* 2001;13:719–33. [PubMed: 11305899]
16. Gaillard WD, Balsamo L, Xu B, et al. fMRI language task panel improves determination of language dominance. *Neurology* 2004;63:1403–8. [PubMed: 15505156]
17. Lehéryc S, Cohen L, Bazin B, et al. Functional MR evaluation of temporal and frontal language dominance compared with the Wada test. *Neurology* 2000;54:1625–33. [PubMed: 10762504]
18. Mahdavi A, Houshmand S, Oghabian MA, et al. Developing optimized fMRI protocol for clinical use: Comparison of different language paradigms. *J Magn Reson Imaging* 2011;34:413–9. [PubMed: 21780233]
19. Partovi S, Konrad F, Karimi S, et al. Effects of covert and overt paradigms in clinical language fMRI. *Acad Radiol* 2012;19:518–25. [PubMed: 22281389]
20. Black DF, Delone DR, Kaufmann TJ, et al. Retrospective analysis of interobserver spatial variability in the localization of Broca's and Wernicke's areas using three different fMRI language paradigms. *J Neuroimaging* 2015;25:626–33. [PubMed: 25496329]
21. Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* 1971;9:97–113. [PubMed: 5146491]
22. Ashtari M, Perrine K, Elbaz R, et al. Mapping the functional anatomy of sentence comprehension and application to presurgical evaluation of patients with brain tumor. *AJNR Am J Neuroradiol* 2005;26:1461–8. [PubMed: 15956516]
23. Tie Y, Whalen S, Suarez RO, Golby AJ. Group independent component analysis of language fMRI from word generation tasks. *Neuroimage* 2008;42:1214–25. [PubMed: 18621548]
24. Da Costa S, van der Zwaag W, Marques JP, et al. Human primary auditory cortex follows the shape of Heschl's gyrus. *J Neurosci* 2011;31:14067–75. [PubMed: 21976491]
25. Friston KJ, Holmes AP, Worsley KJ, et al. Statistical parametric maps in functional imaging: a general linear approach. 1995;2:189–210.
26. ACR–ASNR–SPR Practice parameter for the performance of magnetic resonance imaging (MRI) of the head and neck. Available at <https://www.acr.org/-/media/ACR/Files/Practice-Parameters/mr-headneck.pdf?la=en>. Accessed September 27, 2018.
27. Broca P Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie. *Bull Soc Anat Paris* 1861;6:330–57.

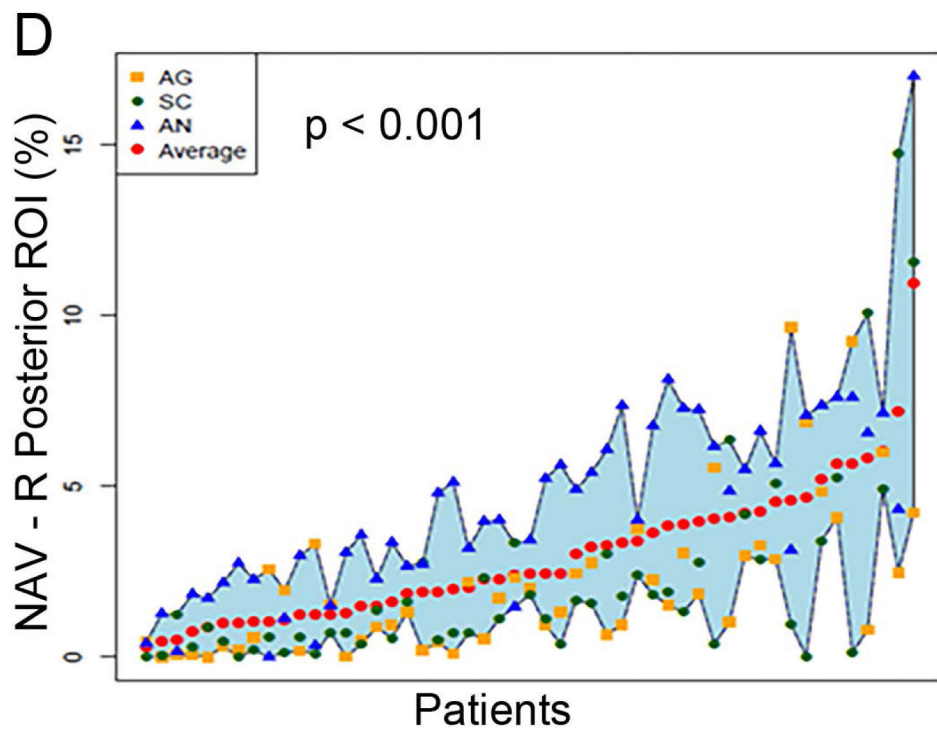
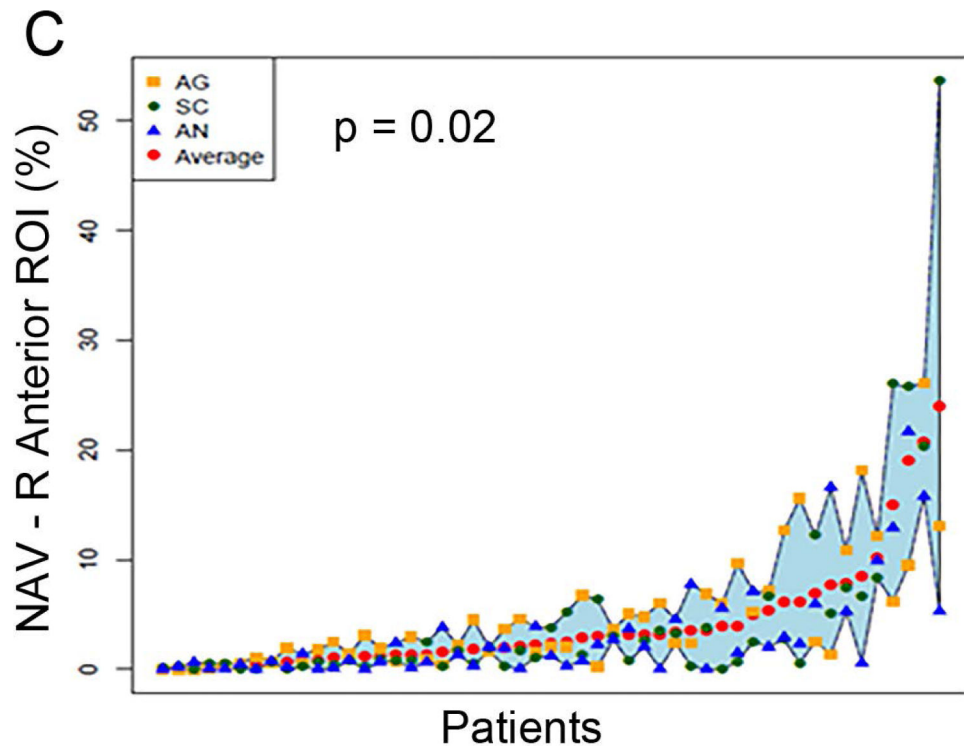
28. Wernicke C Der Aphasische Symptomencomplex Eine Psychologische Studie Auf Anatomischer Basis. Breslau: Cohn & Weigert; 1874.
29. Ross ED The Aprosodias: Functional-anatomic organization of the affective components of language in the right hemisphere. *Arch Neurol* 1981;38:561–9. [PubMed: 7271534]
30. Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. Stuttgart, Ger: Theime; 1988.
31. Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 2003;19:1233–9. [PubMed: 12880848]
32. Tie Y, Suarez RO, Whalen S, et al. Comparison of blocked and event-related fMRI designs for presurgical language mapping. *Neuroimage* 2009;47:T107–15. [PubMed: 19101639]
33. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging* 2012;30:1323–41. [PubMed: 22770690]
34. Deblaere K, Boon PA, Vandemaele P, et al. MRI language dominance assessment in epilepsy patients at 1.0 T: region of interest analysis and comparison with intracarotid amytal testing. *Neuroradiology* 2004;46:413–20. [PubMed: 15127167]
35. Rombouts SARB, Barkhof F, Hoogenraad FG, et al. Test-retest analysis with functional MR of the activated area in the human visual cortex. *AJNR Am J Neuroradiol* 1997;18:1317–22. [PubMed: 9282862]
36. Friedman M The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J Am Stat Assoc* 1937;32:675–701.
37. R Core Team. R: A language and environment for statistical computing. 2017 Available at [www.R-project.org](http://www.R-project.org). Accessed September 27, 2018.
38. McNemar Q Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947;12:153–7. [PubMed: 20254758]
39. Binder JR, Swanson SJ, Hammeke TA, Sabsevitz DS. A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia* 2008;49:1980–97. [PubMed: 18513352]
40. Partovi S, Jacobi B, Rapps N, et al. Clinical standardized fMRI reveals altered language lateralization in patients with brain tumor. *Am J Neuroradiol* 2012;33:2151–7. [PubMed: 22595902]
41. Pak RW, Hadjiabadi DH, Senarathna J, et al. Implications of neurovascular uncoupling in functional magnetic resonance imaging (fMRI) of brain tumors. *J Cereb Blood Flow Metab* 2017;37:3475–87. [PubMed: 28492341]
42. Barnett A, Marty-Dugas J, McAndrews MP. Advantages of sentence-level fMRI language tasks in presurgical language mapping for temporal lobe epilepsy. *Epilepsy Behav* 2014;32:114–20. [PubMed: 24534479]
43. Binder JR, Swanson SJ, Hammeke TA, et al. Determination of language dominance using functional MRI: a comparison with the Wada test. *Neurology* 1996;46:978–84. [PubMed: 8780076]
44. Suarez RO, Whalen S, Nelson AP, et al. Threshold-independent functional MRI determination of language dominance: A validation study against clinical gold standards. *Epilepsy Behav* 2009;16:288–97. [PubMed: 19733509]

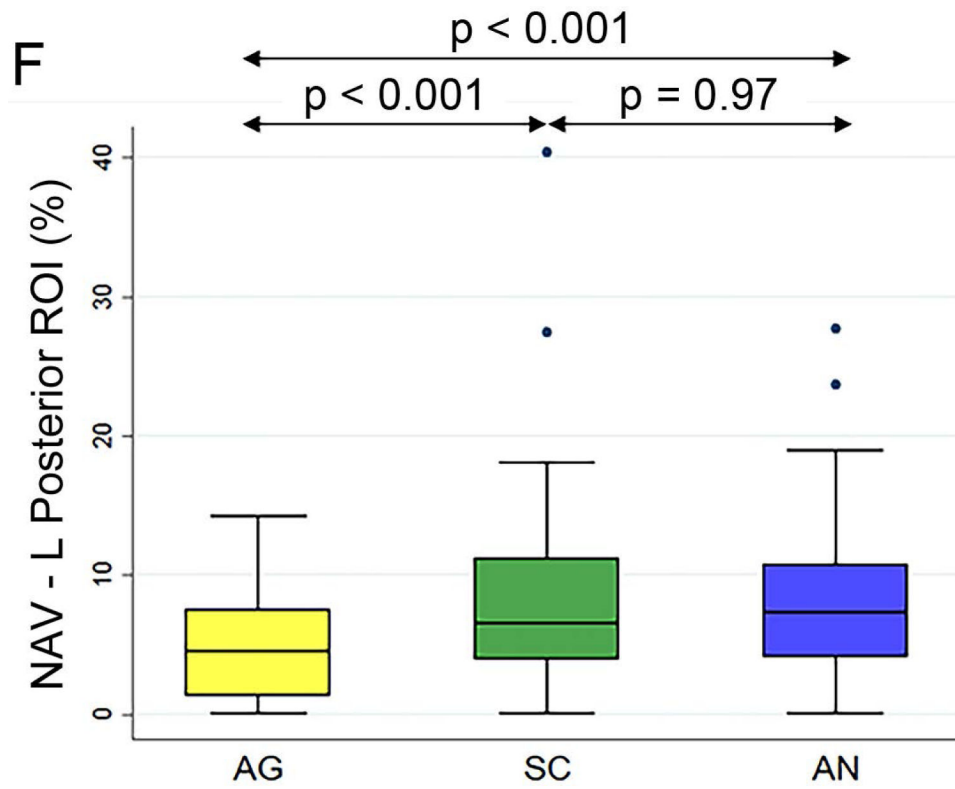
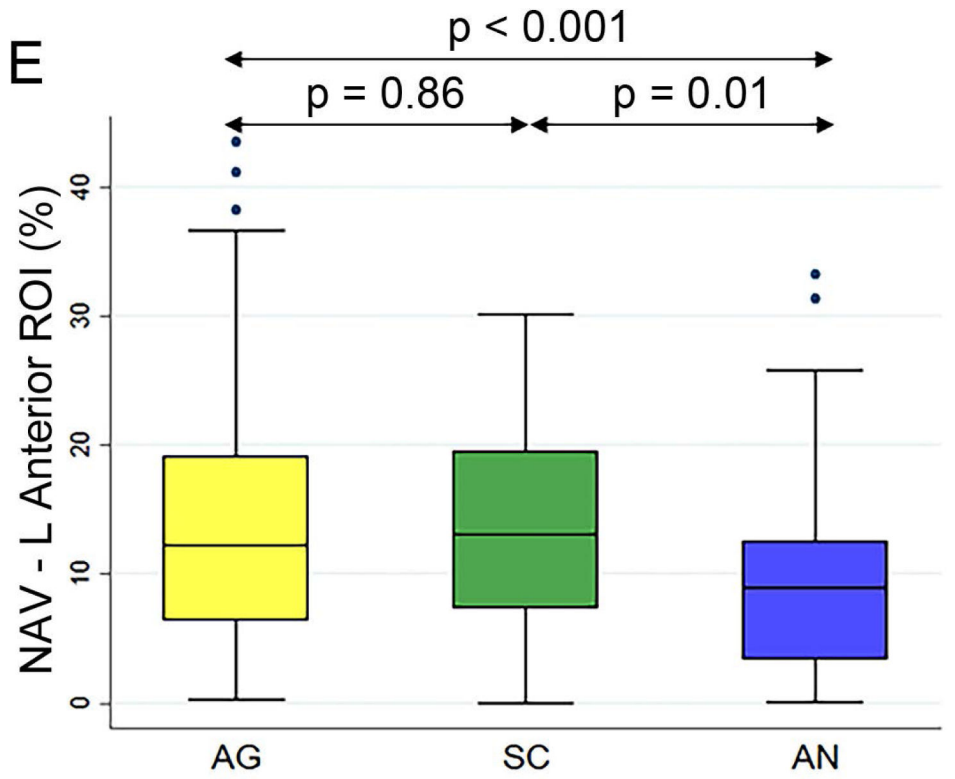


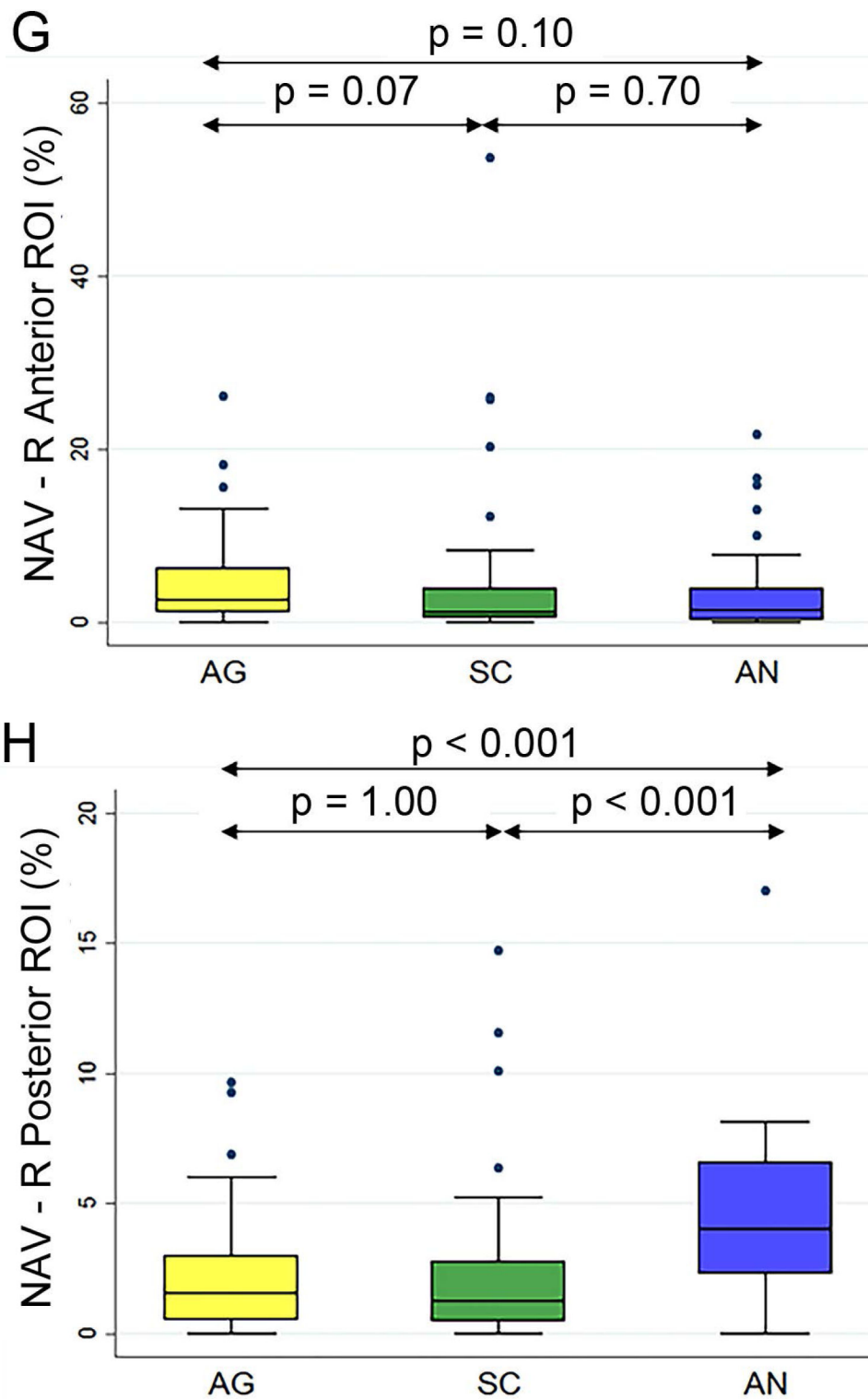
**Figure 1.**

One example patient's language regions-of-interest overlaid onto the structural image. The language regions-of-interest (ROIs) defined in the Montreal Neurological Institute space were reversed-normalized to this patient's own brain space. The anterior ROIs are shown in red for the left hemisphere and green for the right hemisphere. The posterior ROIs are shown in yellow for the left hemisphere and blue for the right hemisphere.









**Figure 2.** The normalized activation volumes of all patients' individual tasks' data and the box plots of the three tasks' normalized activation volumes within the four language regions-of-interest. (A to D) For each region-of-interest (ROI), the normalized activation volumes (NAVs)



averaged across the three tasks (red dots) for each patient are plotted and arranged from the lowest to the highest NAV. Please note the difference in the y-axis scales for the four regions. (E to H) The box plots of the three tasks' NAVs within each ROI. AG: antonym generation task; SC: sentence completion task; AN: auditory naming task; L: left; R: right.

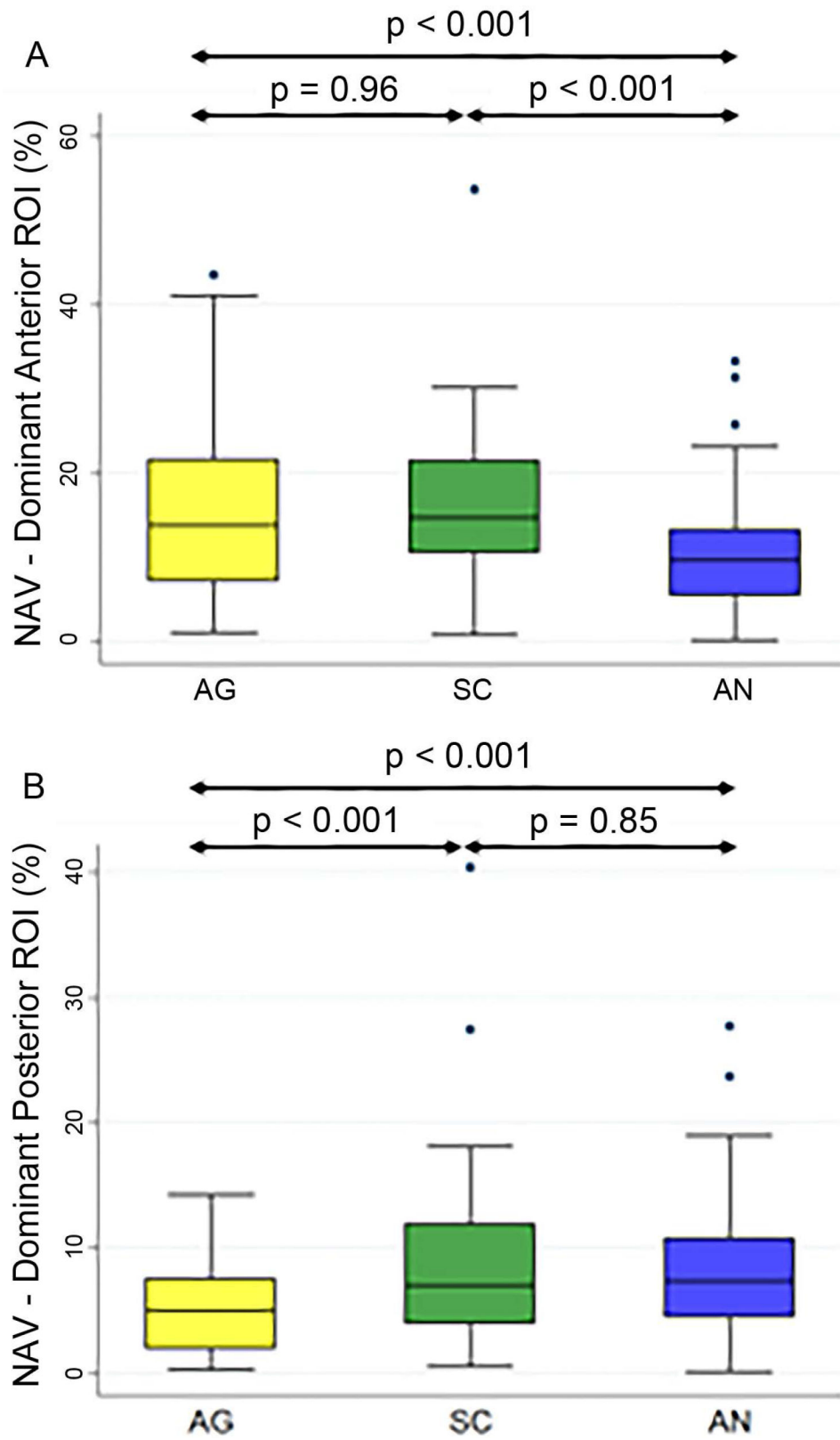
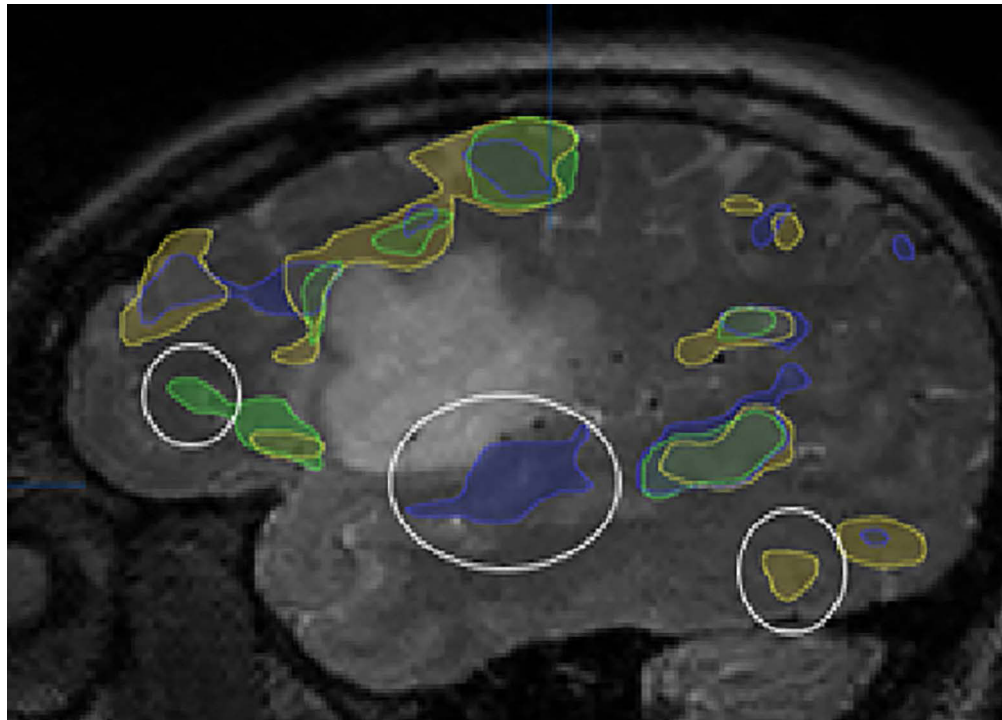
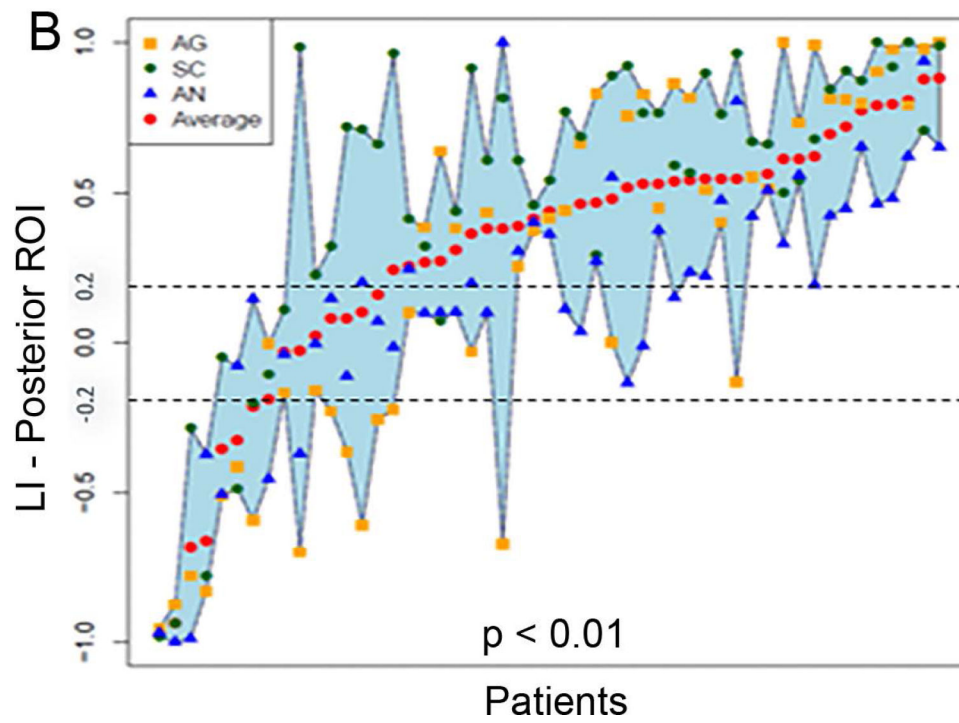
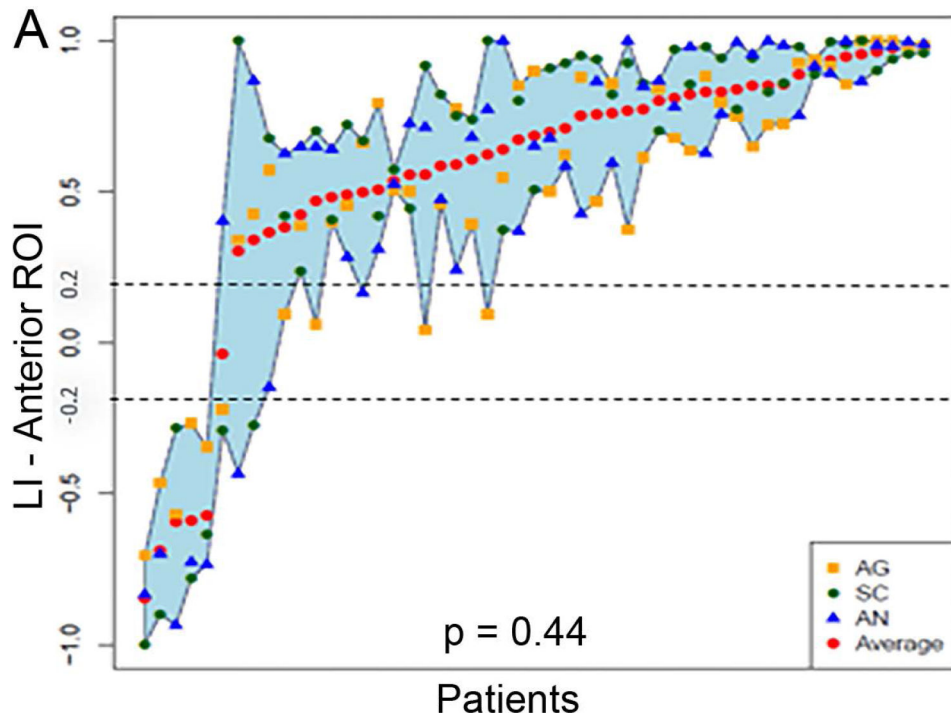


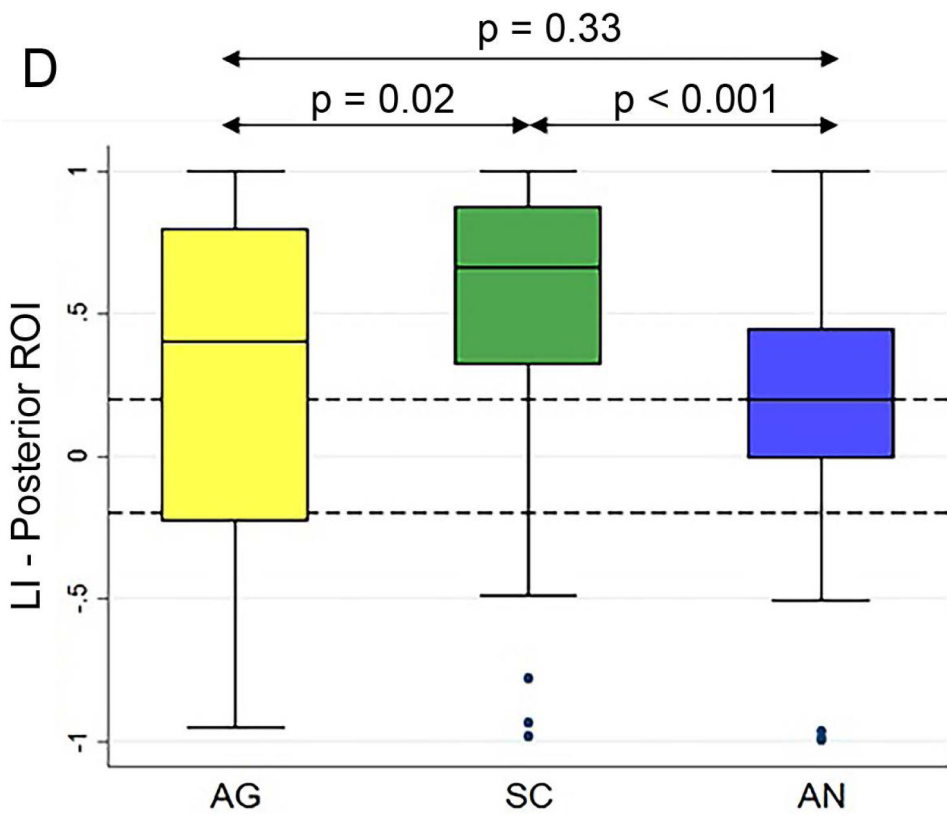
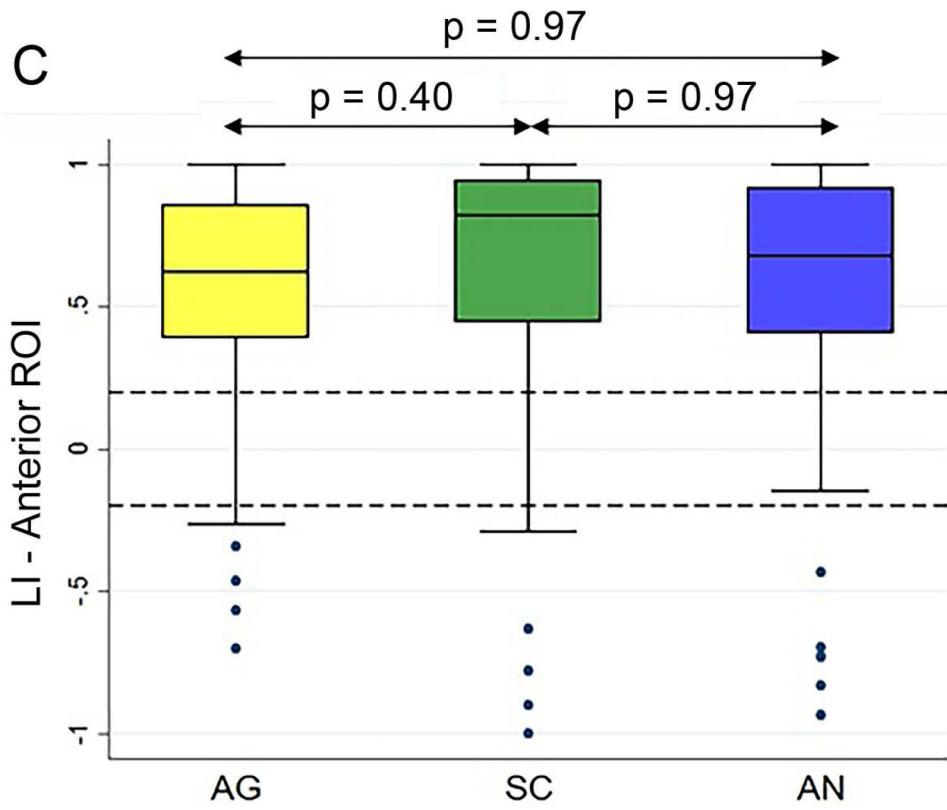
Figure 3.

Box plots of the three tasks' normalized activation volumes within the dominant anterior and posterior language regions-of-interest. NAV: normalized activation volumes; ROI: region-of-interest; AG: antonym generation task; SC: sentence completion task; AN: auditory naming task.



**Figure 4.** Thresholded fMRI maps of the three language tasks of an example patient with a low-grade glioma. The language maps derived from the three tasks are overlaid on the T2-weighted image using 3D Slicer. Results demonstrate the recruitment of cortical structures shared by all three tasks (AG in yellow, SC in green, and AN in blue). The white circles show the activations exclusively elicited by one language task. AG: antonym generation task; SC: sentence completion task; AN: auditory naming task.





**Figure 5.**

Laterality index results of all patients' individual tasks' data and the box plots of the three tasks' LIs within the anterior and posterior language regions-of-interest. (A and B) For the anterior and posterior language regions-of-interest (ROIs), the laterality index (LI) averaged across the three tasks (red dots) for each patient are plotted and arranged from the lowest to the highest LI. (C and D) The box plots of the three tasks' LIs within the anterior and posterior language ROIs. The dotted lines represent the LI cut-offs for hemispheric lateralization for language function ( $LI > 0.2$ , left lateralized;  $-0.2 < LI < 0.2$ , bilateral;  $LI < -0.2$ , right lateralized). AG: antonym generation task; SC: sentence completion task; AN: auditory naming task.

**Table 1.**

Paradigm design of the three language tasks.

<b>Paradigm</b>	<b>Time</b>	<b>Stimulus Modality</b>	<b>Task Block (20 seconds per block)</b>	<b>Baseline (20 seconds per block)</b>
Sentence Completion	4 mins	Visual	Six blocks 4 sentences per block Example: "A square has four___."	Six blocks 4 non-pronounceable non-word letter strings per block Example: "Xbg rkc vhrmgr mbgn sx mhp _____."
Antonym Generation	5 mins	Visual	Seven blocks 5 words per block Example: "Summer"	Eight blocks "+."
Auditory Naming	4 mins	Auditory	Six blocks 4 questions per block Example: "What color is the sky?"	Six blocks Bitonal beeping

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Table 2.**

Pairwise comparison results of the normalized activation volumes within the four language regions-of-interest.

	<b>Task Pairs</b>	<b>p Value (Bonferroni-corrected)</b>	<b>Effect Size</b>
<b>Left Anterior ROI</b>	AG vs. AN	$p < 0.001$	0.63
	SC vs. AN	$p = 0.01$	0.58
	AG vs. SC	$p = 0.86$	0.09
<b>Left Posterior ROI</b>	SC vs. AG	$p < 0.001$	0.56
	AN vs. AG	$p < 0.001$	0.66
	SC vs. AN	$p = 0.97$	0.01
<b>Right Anterior ROI</b>	AG vs. SC	$p = 0.07$	0.02
	AG vs. AN	$p = 0.10$	0.27
	SC vs. AN	$p = 0.70$	0.17
<b>Right Posterior ROI</b>	AN vs. AG	$p < 0.001$	0.80
	AN vs. SC	$p < 0.001$	0.76
	AG vs. SC	$p = 1.00$	0.01

ROI: region-of-interest; AG: antonym generation task; SC: sentence completion task; AN: auditory naming task.

**Table 3.**

Pairwise comparison results of the normalized activation volumes within the language region-of-interest in the dominant hemisphere.

	Task Pairs	p Value (Bonferroni-corrected)	Effect Size
<b>Dominant Anterior ROI</b>	AG vs. AN	$p < 0.001$	0.54
	SC vs. AN	$p < 0.001$	0.60
	AG vs. SC	$p = 0.96$	0.10
<b>Dominant Posterior ROI</b>	SC vs. AG	$p < 0.001$	0.61
	AN vs. AG	$p < 0.001$	0.67
	SC vs. AN	$p = 0.85$	0.07

ROI: region-of-interest; AG: antonym generation task; SC: sentence completion task; AN: auditory naming task.

**Table 4.**

Dice coefficient results within the language regions-of-interest in the dominant hemisphere.

Task Pair	Dice Coefficients (mean $\pm$ standard deviation)	
	Dominant Anterior ROI (n = 43)	Dominant Posterior ROI (n = 26)
AG & SC	0.452 $\pm$ 0.170	0.417 $\pm$ 0.218
AG & AN	0.385 $\pm$ 0.194	0.351 $\pm$ 0.205
SC & AN	0.410 $\pm$ 0.201	0.458 $\pm$ 0.213

ROI: region-of-interest; AG: antonym generation task; SC: sentence completion task; AN: auditory naming task; n: number of patients.

**Table 5.**

Pairwise comparison results of the laterality index within the language region-of-interest in the dominant hemisphere.

	Task Pairs	p Value (Bonferroni-corrected)	Effect Size
<b>Dominant Anterior ROI</b>	AG vs. AN	p = 0.97	0.08
	SC vs. AN	p = 0.97	0.17
	AG vs. SC	p = 0.40	0.30
<b>Dominant Posterior ROI</b>	SC vs. AG	p = 0.02	0.53
	SC vs. AN	p < 0.001	0.88
	AG vs. AN	p = 0.33	0.10

ROI: region-of-interest; AG: antonym generation task; SC: sentence completion task; AN: auditory naming task.

**Table 6.**

Number of patients with discordant hemispheric lateralization within the language regions-of-interest.

	<b>Anterior ROI</b>	<b>Posterior ROI</b>
Discordance among all three tasks	0	4
AG not in agreement with the other two tasks	3	6
SC not in agreement with the other two tasks	1	3
AN not in agreement with the other two tasks	4	12
Total	8	25

ROI: region-of-interest; AG: antonym generation task; SC: sentence completion task; AN: auditory naming task.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript