NeuroImage: Clinical 14 (2017) 552–565

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Damage to white matter bottlenecks contributes to language impairments after left hemispheric stroke

Joseph C. Griffis[a,*], Rodolphe Nenert[b], Jane B. Allendorfer[b], Jerzy P. Szaflarski[b]

[a] University of Alabama at Birmingham, Department of Psychology, United States
[b] University of Alabama at Birmingham, Department of Neurology, United States

ARTICLE INFO

Article history:
Received 4 November 2016
Received in revised form 16 February 2017
Accepted 23 February 2017
Available online 24 February 2017

Keywords:
Aphasia
Stroke
White matter
Language
MRI

ABSTRACT

Damage to the white matter underlying the left posterior temporal lobe leads to deficits in multiple language functions. The posterior temporal white matter may correspond to a bottleneck where both dorsal and ventral language pathways are vulnerable to simultaneous damage. Damage to a second putative white matter bottleneck in the left deep prefrontal white matter involving projections associated with ventral language pathways and thalamo-cortical projections has recently been proposed as a source of semantic deficits after stroke. Here, we first used white matter atlases to identify the previously described white matter bottlenecks in the posterior temporal and deep prefrontal white matter. We then assessed the effects of damage to each region on measures of verbal fluency, picture naming, and auditory semantic decision-making in 43 chronic left hemispheric stroke patients. Damage to the posterior temporal bottleneck predicted deficits on all tasks, while damage to the anterior bottleneck only significantly predicted deficits in verbal fluency. Importantly, the effects of damage to the bottleneck regions were not attributable to lesion volume, lesion loads on the tracts traversing the bottlenecks, or damage to nearby cortical language areas. Multivariate lesion-symptom mapping revealed additional lesion predictors of deficits. Post-hoc fiber tracking of the peak white matter lesion predictors using a publicly available tractography atlas revealed evidence consistent with the results of the bottleneck analyses. Together, our results provide support for the proposal that spatially specific white matter damage affecting bottleneck regions, particularly in the posterior temporal lobe, contributes to chronic language deficits after left hemispheric stroke. This may reflect the simultaneous disruption of signaling in dorsal and ventral language processing streams.

© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Damage to long-range white matter (WM) pathways likely contributes substantially to language deficits after left hemisphere stroke. Damage to the WM underlying the left posterior superior and middle temporal gyri (pSTG and pMTG) has been consistently implicated as a source of deficits in multiple language domains including comprehension (Dronkers et al., 2004; Geva et al., 2012; Henseler et al., 2014; Pustina et al., 2016; Yourganov et al., 2016), naming (Baldo et al., 2013; Harvey and Schnur, 2015; Henseler et al., 2014; Pustina et al., 2016; Yourganov et al., 2016), repetition (Butler et al., 2014; Henseler et al., 2014; Pustina et al., 2016; Yourganov et al., 2016), and phonology (Butler et al., 2014). The presence of lesions affecting the posterior temporal WM and disrupting posterior temporal connectivity also predicts poor responses to language therapies (Bonilha et al., 2015; Fridriksson, 2010).

Why might damage to the WM in this area have such broadly negative impacts on language outcomes? Portions of the WM under the pSTG/pMTG contain projections associated with multiple long-range fiber pathways (Turken and Dronkers, 2011), including dorsal (sensori-motor) and ventral (associative) language pathways (Kümmerer et al., 2013; Saur et al., 2008). Fibers associated with at least three language-relevant tracts traverse this region – the arcuate fasciculus (AF – dorsal stream), inferior longitudinal fasciculus (ILF – ventral stream), and inferior fronto-occipital fasciculus (IFOF – ventral stream) (Catani and Mesulam, 2008; Turken and Dronkers, 2011). Thus, it has been proposed that the WM in this area corresponds to a structural weak point, or "bottleneck", where multiple language-relevant pathways are vulnerable to simultaneous disruption by focal damage (Turken and Dronkers, 2011). The observation that fibers associated with the anterior thalamic radiations (ATR – thalamo-cortical), uncinate fasciculus (UF – ventral stream), and inferior fronto-occipital fasciculus (IFOF – ventral stream) form a bottleneck in the prefrontal WM near areas where damage is associated with chronic deficits in semantic recognition supports the proposal that damage to bottleneck regions may

* Corresponding author at: Department of Neurology and UABEC, University of Alabama at Birmingham, 312 Civitan International Research Center, 1719, 6th Avenue South, Birmingham, AL 35294-0021, United States.
E-mail address: joegriff@uab.edu (J.C. Griffis).

http://dx.doi.org/10.1016/j.nicl.2017.02.019
2213-1582/© 2017 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

play a role in chronic language deficits after stroke (Mirman et al., 2015a). Indeed, this proposal is in accord with recent evidence suggesting that lesions affecting areas of high tract overlap are associated with post-stroke deficits in multiple cognitive domains (Corbetta et al., 2015).

However, the conclusions that can be drawn from previous studies linking bottleneck lesions to language deficits are limited because the bottleneck regions were identified as part of *post-hoc* exploratory analyses based on the results of voxel-wise lesion-symptom mapping (Dronkers et al., 2004; Mirman et al., 2015a; Turken and Dronkers, 2011). The effects of damage to *a priori* identified bottlenecks in these regions on language outcomes have not been investigated. We aimed to bridge this gap by characterizing how deficits in measures of verbal fluency, picture naming, and auditory semantic processing relate to lesions affecting the bottleneck regions described by previous reports. We expected that damage to the bottleneck underlying the left pSTG/pMTG would be associated with chronic impairments on all language measures, as broad deficits would be expected to follow the simultaneous disruption of both ventral and dorsal language streams. Based on the report by Mirman et al. (2015a), we expected that damage to the prefrontal bottleneck might be associated with deficits in picture naming and category judgments, which could influence performance on the other tasks. To enable stronger conclusions about our specific findings, we demonstrate that the effects of damage to these bottleneck regions are not attributable to lesion loads on the tracts traversing them or to concomitant cortical damage. In a second analysis using a data-driven approach, we thoroughly characterize the lesion-deficit relationships in these patients using multivariate lesion-symptom mapping with lesion volume control (Zhang et al., 2014). Exploratory fiber tracking was also performed to further characterize potential lesion effects on inter-regional connections.

2. Methods

2.1. Participants

All procedures were approved by the Institutional Review Boards of the participating institutions and performed in accordance with Declaration of Helsinki ethics principles and principles of informed consent. Data were collected from 43 patients with chronic left hemispheric stroke participating in different studies by our laboratory. Patients were excluded if they had diagnoses of degenerative/metabolic disorders, diagnoses of severe depression or other psychiatric disorders, were pregnant, were not fluent in English, or had any contraindication to MRI/fMRI. All patients had a single left hemispheric stroke and received a clinical diagnosis of aphasia following the initial insult, which occurred at least 1 year prior to data collection. Current scores on clinical aphasia measures (i.e. WAB/BDAE) were not available for all patients, but some patients were likely not aphasic at the time of data collection as indicated by their performance on the language tasks (i.e. some patients performed at levels comparable to healthy controls; see Fig. 1). This sample, which features patients with varying degrees of impairment, is well-suited for the lesion-behavior analyses employed in this study. No patients had right hemispheric stroke. The mean patient age was 53 (SD = 15; range = 23–90), 25 patients were male, and the mean pre-stroke handedness as determined by the Edinburgh Handedness Inventory (Oldfield, 1971) was 0.85 (SD = 0.43; range = −1.0–1.0). Data on educational background were not available for these patients. Behavioral data collected from a group of 43 age/sex/handedness-matched healthy controls were included as a reference. The mean control age was 54 (SD = 15; range = 19–74), 23 controls were male, and the mean control EHI was 0.80 (SD = 0.41; range = −0.9–1.0). A detailed

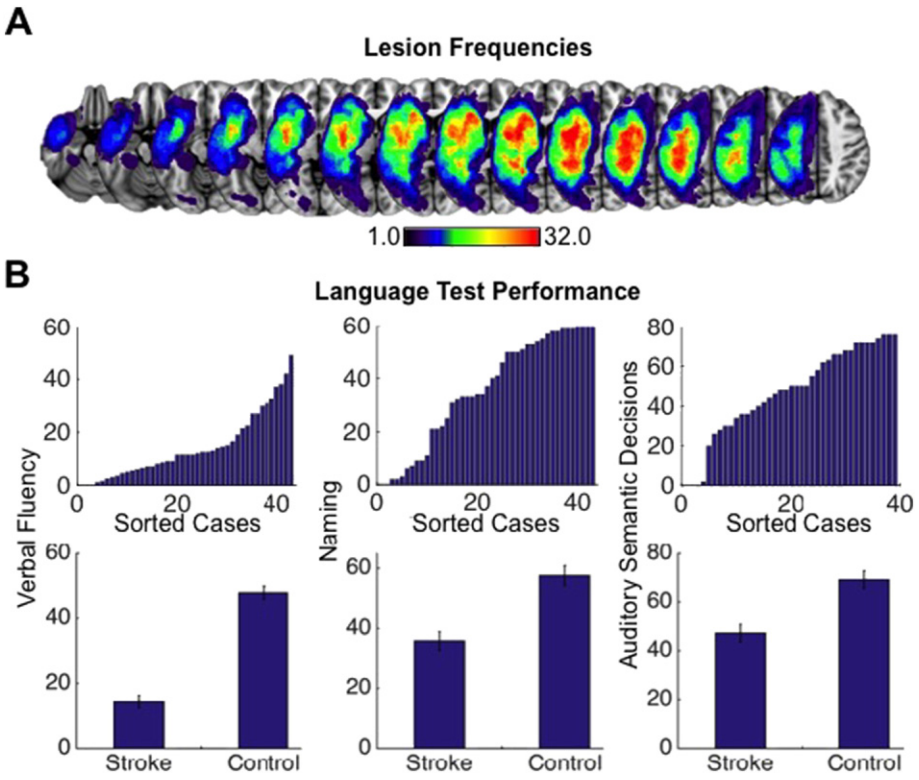

Fig. 1. Characterization of lesion and language test data. **A**. Lesion frequency overlaps are shown on axial slices from a template brain. Color bar values range from 1 to 32, and indicate the number of patients with lesions at each voxel. **B**. Plots on the top row show sorted scores (from low to high) for the Verbal Fluency (left), Naming (middle), and Auditory Semantic Decision (right) from the stroke patients. Bar graphs on the bottom row show mean scores for stroke patients and for 43 age/sex/handedness-matched healthy controls (error bars indicate standard error of the mean).

characterization of patient characteristics is provided in Supplementary Table 1.

2.2. Neuroimaging data collection

MRI Data were acquired at the University of Alabama at Birmingham using a 3T head-only Siemens Magnetom Allegra scanner provided by the Civitan International Research Center Functional Imaging Laboratory. These data consisted of 3D high-resolution T1-weighted anatomical scans (TR/TE = 2.3 s/2.17 ms, FOV = 25.6 × 25.6 × 19.2 cm, matrix = 256 × 256, flip angle = 9 degrees, slice thickness = 1 mm). MRI data were also collected at the Cincinnati Children's Hospital Medical Center on a 3T research-dedicated Philips MRI system provided by the Imaging Research Center. These data consisted of 3D high-resolution T1-weighted anatomical scans (TR/TE = 8.1 s/2.17 ms, FOV = 25.0 × 21.0 × 18.0 cm, matrix = 252 × 211, flip angle = 8 degrees, slice thickness = 1 mm).

2.3. Lesion identification

MRI data were processed using Statistical Parametric Mapping (SPM) (Friston et al., 1995) version 12 running in MATLAB r2014b (The MathWorks, Natick MA, USA). T1-weighted scans were segmented and normalized using the unified normalization procedure implemented in SPM12. Lesion probability maps were created for each patient using a probabilistic lesion classification algorithm implemented in the *lesion_gnb* toolbox for SPM12 (Griffis et al., 2016a). Lesion probability maps were manually thresholded to ensure that the resulting binary masks precisely reflected the lesions, and then resampled to 2 mm isotropic resolution. Lesion frequencies are shown in Fig. 1A.

2.4. Language measures

All patients completed a battery of language assessments that included the Boston Naming Test (BNT) (Kaplan et al., 2001), Semantic Fluency Test (SFT) form A (Kozora and Cullum, 1995), and Controlled Oral Word Association Test (COWAT) form A (Lezak et al., 1995) prior to MRI scanning. The BNT is a picture naming test that consists of black and white line drawings of animate and inanimate items. Naming abilities were measured as the number of correctly named pictures. The SFT and COWAT both involve generating words in response to a given prompt within a 1 min time limit. The SFT is a measure of verbal semantic fluency that uses semantic category prompts (animals, fruits/vegetables, things that are hot), whereas the COWAT is a measure of verbal letter fluency that uses letter category prompts (C, F, and L). These measures and analogous measures have previously been employed as measures of language function by other anatomical (Almairac et al., 2014; Baldo et al., 2006; Baldo et al., 2013; Bizzi et al., 2012; Hope et al., 2015; Thames et al., 2012) and functional neuroimaging (Blasi et al., 2002; Griffis et al., 2016b; Peck et al., 2004; Perani et al., 2003; Szaflarski et al., 2013) studies investigating language deficits in clinical populations.

As part of a separate fMRI scan, an auditory semantic decision task was completed in the scanner, and the behavioral data were used as a measure of auditory semantic processing. This task robustly activates networks involved in semantic comprehension (Binder et al., 1997; Binder et al., 2008; Frost et al., 1999; Szaflarski et al., 2002) even in patients with chronic stroke (Eaton et al., 2008; Griffis et al., 2016b; Szaflarski et al., 2013). We note that the fMRI data associated with this task are characterized separately (Griffis et al., 2016b), as this study is focused on structural data only. Participants heard 50 spoken English nouns designating different animals, and decided if the animals were both “native to the United States” and “commonly used by humans” (e.g. for food, clothing, or labor), as indicated by button press. Thus, this task requires patients to make a conjunction decision about whether each animal meets both criteria. A mock run with five trials outside of the scanner confirmed that patients understood the task. Auditory semantic decision task data were missing for 4 patients due to hardware issues.

Scores on the SFT and COWAT were strongly correlated ($r = 0.92$), and patients' general verbal fluency was measured as the average number of words generated for both the SFT and COWAT. A similar composite measure of semantic and letter fluency outcomes was used by a recent study investigating the effects of tract disconnection on language abilities in chronic stroke patients (Hope et al., 2015). However, we note that there is likely only partial overlap between the neural substrates of these processes (Biesbroek et al., 2015; Chouiter et al., 2016; Costafreda et al., 2006; Heim et al., 2009), and there is also evidence for some differences in the contributions of temporal/parietal and frontal areas to semantic and letter fluency, respectively (Baldo et al., 2006; Biesbroek et al., 2015; Chouiter et al., 2016). The use of an average verbal fluency measure may, therefore, make our analyses less sensitive to regions that do not support both processes, although separate analyses may also have low sensitivity to such regions due to the high correlation between the two tasks. Nonetheless, lesion effects were characterized separately for each verbal fluency task in the Supplementary Material (Supplementary Analyses 1 and 2).

We also note that despite having a language component, verbal fluency tasks are often regarded as primarily measuring high-level executive function (Alvarez and Emory, 2006; Jurado and Rosselli, 2007). However, recent evidence indicates that verbal fluency performance may relate more to language ability than to general executive function (Whiteside et al., 2015), and suggests that post-stroke deficits in semantic fluency are strongly related to impaired lexical retrieval processes (Bose et al., 2016). This is perhaps not surprising, as functional neuroimaging (Costafreda et al., 2006; Heim et al., 2009; Vitali et al., 2005) and lesion (Baldo et al., 2006; Hope et al., 2015) studies have implicated classical language regions and connections (e.g. left inferior frontal gyrus – IFG, AF, etc.) in supporting verbal fluency. Even so, the complex nature of these tasks must be considered in interpreting the results.

Scores on the BNT ($r = 0.76$) and semantic decision task ($r = 0.53$) showed positive correlations to scores on the composite verbal fluency measure. Scores on the BNT were also positively correlated with scores on the auditory semantic decision task ($r = 0.43$). Language test scores for each patient, along with mean scores for both patients and healthy controls are shown in Fig. 1B.

2.5. White matter bottleneck ROIs

We defined two a priori regions of interest (ROI) corresponding to the putative bottleneck regions described by previous studies (Mirman et al., 2015a; Mirman et al., 2015b; Turken and Dronkers, 2011). We first defined an anterior bottleneck ROI according to precisely the same procedure used by Mirman et al. (2015a,b) to identify the bottleneck region described in their reports. This procedure consisted of thresholding probabilistic atlas labels (ICBM-DTI atlas – available at http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases (Oishi et al., 2008)) for the left UF, IFOF, and ATR at a 20% tract probability threshold, and intersecting the thresholded labels. We note that the 20% tract probability threshold was chosen because this was the threshold used by a previous study that identified the anterior bottleneck region (Mirman et al., 2015a). Voxels contained within the intersection of these tract labels were defined as the anterior bottleneck ROI. Next, we defined a posterior bottleneck ROI according to the same procedure, but using labels for three of the language-relevant tracts previously identified in the posterior temporal WM: the left ILF, IFOF, and SLF/AF (Turken and Dronkers, 2011). The same (i.e. 20%) tract probability threshold was used to ensure consistency between the ROI definitions. To emphasize, the tracts used to define the posterior bottleneck region were chosen based on previous reports that fibers associated with these tracts traverse portions of the WM underlying the left pMTG/pSTG where lesions are

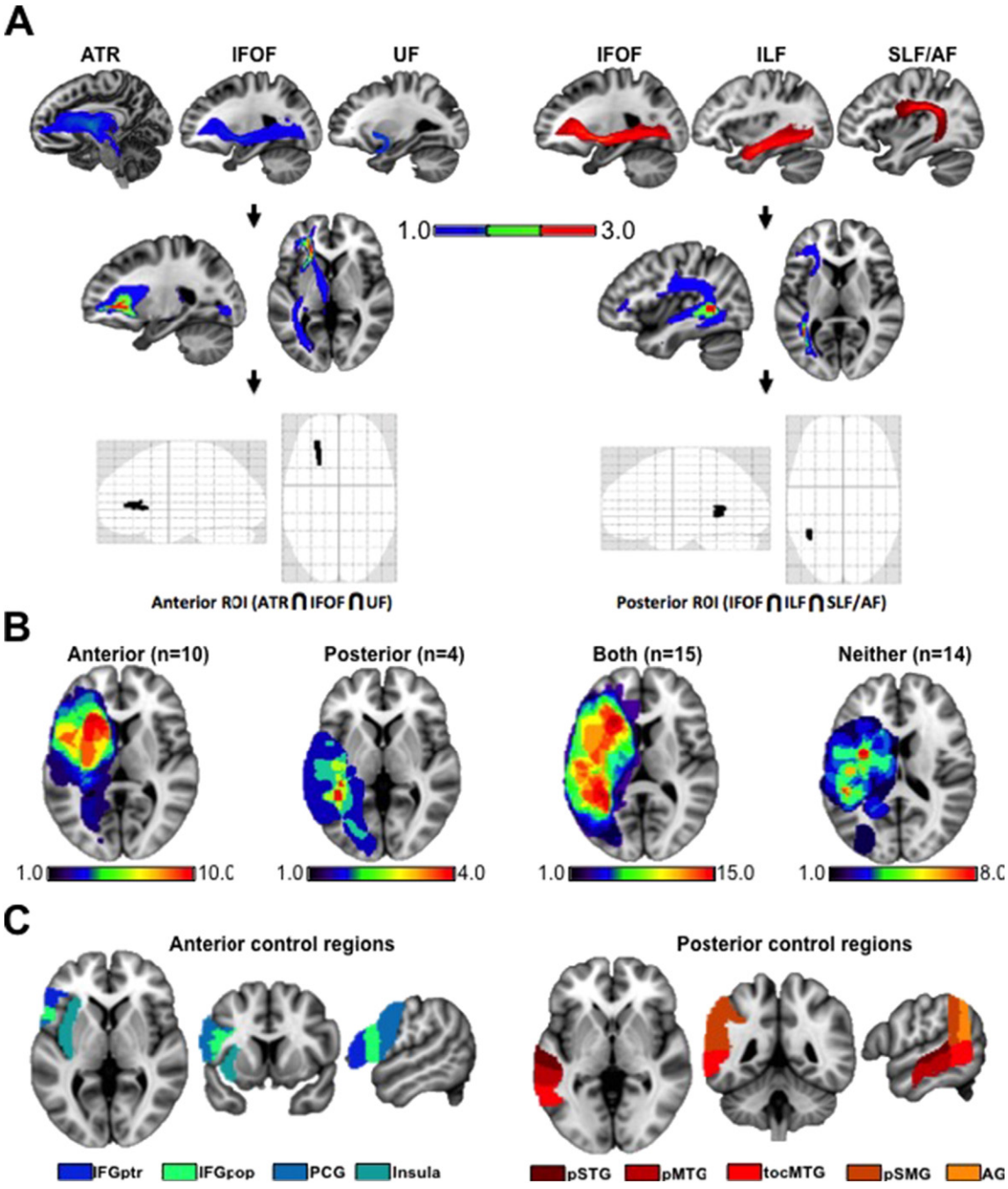

Fig. 2. Regions of interest. **A**. The process for defining each ROI is illustrated in three steps. First, a threshold of 20% was applied to the atlas labels corresponding to the left hemispheric tracts expected to pass through the anterior (top – blue) and posterior (top – red) bottleneck ROIs. Next, the thresholded tracts were binzarized and overlapped (middle – color bar values indicate number of tracts at each voxel). Finally, the anterior bottleneck ROI was defined as the intersection of the voxels contained in the thresholded ATR, IFOF, and UF labels (bottom left), and the posterior bottleneck ROI was defined as the intersection of the voxels contained in the thresholded IFOF, ILF, and SLF/AF labels (bottom right). **B**. Representative slices overlaid with lesion-frequency maps for patients with damage to only the anterior bottleneck ROI (left), only the posterior bottleneck ROI (left middle), both bottleneck ROIs (right middle), and neither bottleneck ROI (right) are shown. **C**. Harvard-Oxford cortical atlas labels for the left inferior frontal gyrus pars triangularis (IFGptr) and pars opercularis (IFGpop), precentral gyrus (PCG), and insula (left) were used to define the cortical control regions for the anterior bottleneck ROI (left). Labels for the left posterior superior and middle temporal gyri (pSTG/pMTG), temporo-occipital middle temporal gyrus (tocMTG), posterior supramarginal gyrus (pSMG), and angular gyrus (AG) were used to define the cortical control regions for the posterior bottleneck ROI (right). Cortical labels were thresholded to contain only voxels with grey matter tissue probabilities >20%. Note: ∩ indicates the intersection operation.

associated with impaired comprehension (Turken and Dronkers, 2011). While additional tracts (i.e. MDLF and tapetum) were also reported to traverse the white matter near this region (Turken and Dronkers, 2011), these regions were not included in the ROI definition because probabilistic labels for these tracts were not available in the atlas, and their importance for language is not clear. Both ROIs were resampled to 2 mm isotropic resolution. After resampling, the posterior ROI contained 47 voxels, and the anterior ROI contained 66 voxels. The process used to define each ROI is outlined in Fig. 2A.

2.6. Cortical control ROIs

Additional cortical ROIs were defined for use as control variables in subsequent hierarchical regression analyses (described in Section 2.7).

These cortical ROIs were defined using the Harvard-Oxford maximum probability cortical atlas (Fischl et al., 2004) and were intended to account for damage to language-relevant areas near each bottleneck ROI (Binder et al., 2009; Friederici and Gierhan, 2013; Price, 2010; Price, 2012; Vigneau et al., 2006).

Cortical control ROIs for the anterior bottleneck ROI were defined as the left IFG pars triangularis (IFGptr), IFG pars opercularis (IFGpOp), pre-central gyrus (PCG), and insula. Broadly speaking, these regions are commonly implicated in the control and output of speech (Bates et al., 2003; Berker et al., 1986; Fama et al., 2017; Price, 2010; Price, 2012; Yourganov et al., 2016), and also in the support of goal-directed cognition and behavior (Dosenbach et al., 2007; Geranmayeh et al., 2014; Swick et al., 2008). Briefly, the IFGpTr/pOp are implicated in high-level language processes such as verbal fluency (Costafreda et al., 2006; Heim et al., 2009; Henseler et al., 2014) and likely support top-down selection/retrieval, sequencing, and motor planning (Binder and Desai, 2011; Blasi et al., 2002; Holland et al., 2011; Price, 2010; Price, 2012; Thompson-Schill et al., 1997), the PCG contains primary motor areas involved in speech motor output (Bizzi et al., 2012; Fama et al., 2017; Fridriksson et al., 2013; Price, 2010; Yourganov et al., 2016) and the insula (particularly the anterior portion) may support the preparatory control of speech motor output (Dronkers, 1996; Price, 2012; Riecker et al., 2005). We note that the atlas used to define these ROIs did not contain insular subdivisions, and so the entire insular label was used to define the insular ROI.

Cortical ROIs to control for the posterior bottleneck ROI were defined as the left pSTG, pMTG, temporo-occipital MTG (tocMTG), posterior supramarginal gyrus (pSMG), and angular gyrus (AG). In general, pMTG/pSTG and temporo-parietal regions such as the AG are implicated in supporting functions relevant to language comprehension (Bates et al., 2003; Dronkers et al., 2004; Henseler et al., 2014; Mirman et al., 2015a) and semantics Binder et al., 2009; Binder and Desai, 2011; Price, 2010; Skipper-Kallal et al., 2015) respectively. However, regions such as the pSTG/pSMG likely contribute to other functions including speech output and repetition (Henseler et al., 2014; Mirman et al., 2015a; Yourganov et al., 2016; Pustina et al., 2016).

All cortical ROIs were masked to include only voxels with a grey matter tissue probability of at least 20% as determined using the grey matter (GM) tissue probability map (TPM) included in SPM12, and are shown in Fig. 2C.

2.7. Multiple regression analyses

Multiple linear regressions assessed whether damage to each bottleneck ROI predicted deficits in task performance. Binary lesion status (lesion vs. no lesion) was used to designate lesion effects at each ROI (rather than a continuous measure such as percent lesion overlap) because each ROI contained only a small number of voxels. 14 patients did not have damage at either ROI, 10 patients had damage to only the anterior ROI, 4 patients had damage to only the posterior ROI, and 15 patients had damage to both ROIs. Representative slices with lesion frequencies for each subgroup of patients are shown in Fig. 2B. As noted in Section 2.4, 4 patients were excluded from the auditory semantic decision analyses due to hardware issues that resulted in a failure to collect behavioral data in-scanner; 1 of these patients did not have damage to either ROI, 1 had damage to both ROIs, and 2 had damage to only the anterior ROI.

We first sought to determine if the lesion status information for the bottleneck regions significantly predicted task performance while controlling for lesion volume. Thus, we fit a 4-predictor model to each task that included (1) lesion status for the posterior bottleneck, (2) lesion status for the anterior bottleneck, (3) the interaction of the lesion status terms (i.e. to account for the possibility that the effect of damage to both regions was not additive), and (4) total lesion volume as predictors. Continuous variables were converted to z-scores so that they were on the same scale. We refer to these models as the "4-predictor models" in the rest of the text. These models were considered significant if the *p*-value for the *F*-test against the intercept-only model survived Bonferroni-Holm correction to control the FWE at 0.05 across all three (i.e. one for each task) models (Aickin and Gensler, 1996). Parameter estimates for the variables in each of these models were considered significant if the *p*-value of the test statistic survived Bonferroni-Holm correction to control the FWE at 0.05 across all 4 variables in each model.

For each task, bottleneck lesion status predictors that were significant in the 4-predictor models were then entered into a second analysis using a hierarchical multiple regression approach. These models were intended to determine if the significant bottleneck predictors identified in the 4-predictor models included information about task performance beyond that attributable to tract and cortical lesion loads. Thus, lesion loads were calculated for (1) the tracts used to define each bottleneck ROI, and (2) the cortical controls for each bottleneck ROI. Lesion loads were calculated as the proportion of voxels that overlapped with each patient's binary lesion mask (while we excluded voxels corresponding to the bottleneck ROIs, this did not affect the outcome of the analyses – lesion load estimates that included these voxels showed strong [r ~ 0.99] correlations to the original estimates). For each task, the hierarchical regression analysis was performed in two steps. In the first step, lesion loads for the tracts used to define each significant bottleneck lesion status predictor from the 4-predictor model for the same task, and (2) lesion loads for the cortical control ROIs defined for each significant bottleneck lesion status predictor from the 4-predictor model for the same task were entered as predictors of task performance. We refer to the models fit in this step as the "lesion load control models". In the second step, the significant bottleneck lesion status information from the 4-predictor model for the same task was added to the model, and the change in R^2 was assessed to determine if the bottleneck lesion status information significantly improved the model fit. We refer to the models including both cortical/tract lesion loads and bottleneck lesion statuses as the "full models". To summarize, in the first block of each hierarchical regression, which we refer to as the "lesion load control models", lesion loads for (1) the tracts used to define the significant bottleneck lesion status predictor(s) from the 4-predictor model, and (2) the cortical control ROIs defined for the significant bottleneck ROI predictor(s) from the 4-predictor model were entered as predictors. In the second block, which we refer to as the "full model", the significant bottleneck ROI predictor(s) from the 4-predictor model were added to the lesion load control model, and the change in model R^2 was evaluated to determine whether the addition of the bottleneck ROI predictor(s) significantly improved the model fit. Changes in model fits between the lesion load control models and the full models were considered significant if the *p*-value for the *F*-test on the R^2 change statistic survived Bonferroni-Holm correction to control the FWE at 0.05 across all three R^2 change tests (one for each task). To illustrate that the conclusions drawn from these analyses are robust against changes in ROI definition, control analyses were performed using ROIs defined using an alternate white matter atlas (Rojkova et al., 2016), and are provided in the Supplementary Material (Supplementary Analysis 4). Additional control analyses provided in the Supplementary Material further illustrate that the addition of lesion loads for other cortical areas (i.e. frontal operculum and anterior SMG) and total lesion volume to the lesion load control models do not significantly improve model fits (Supplementary Analysis 5).

We note that the hierarchical approach used here is conceptually similar to the approach used by Hope et al. (2015) in their comparison of tract lesion load and tract disconnection measures as predictors of chronic language outcomes. Our approach is also more comprehensive than the approaches used by many other studies investigating WM lesion contributions to chronic language deficits, as controls for concomitant tract and cortical damage are not typically employed (Forkel et al., 2014; Geva et al., 2015; Hope et al., 2015; Ivanova et al., 2016; Kümmerer et al., 2013; Marchina et al., 2011).

2.8. Multivariate lesion-symptom mapping with lesion volume control

Multivariate lesion-symptom mapping using support vector regression (SVR-LSM) was performed using the SVR-LSM toolbox (Zhang et al., 2014). Unlike traditional lesion-symptom mapping approaches that separately test the effects of damage to each voxel on behavioral scores, SVR-LSM identifies lesion-behavior relationships at all voxels simultaneously, making it more sensitive for detecting lesion-symptom relationships than traditional approaches (Zhang et al., 2014). To control for lesion volume effects, we utilized the direct total lesion volume control (DTLVC) option in the SVR-LSM toolbox. This method normalizes each patient's lesion map to have a unit norm, such that the lesion status of each voxel is equal to either 0 or the reciprocal of the norm ($1/\sqrt{\text{lesion volume}}$), and provides better control than the regression of behavioral scores on lesion volume (Zhang et al., 2014). Analyses without lesion volume control are provided as supplements (Supplementary Analysis 3). SVR-LSM + DTLVC analyses only considered voxels that were lesioned in at least 10 patients. While there is not a consensus regarding whether SVR-LSM requires multiple comparisons correction (Fama et al., 2017; Mirman et al., 2015b), lesion-behavior relationships that survived a False Discovery Rate (Benjamini and Hochberg, 1995; Genovese et al., 2002) (FDR) correction threshold of 0.10 as determined by 2000 permutation tests are reported here (Mirman et al., 2015b). We note that unlike FWE-controlling procedures, FDR correction estimates the proportion of discoveries that are likely to be false at a given threshold, and an FDR threshold of 0.10 indicates that on average no >10% of significant voxels are expected to be false discoveries (Benjamini and Hochberg, 1995). Additionally, unlike predictive SVR, where a model is trained to predict outcomes for new cases, the SVR-LSM model is intended to identify significant lesion predictors of behavioral outcomes and is fit to the entire dataset (Zhang et al., 2014). We used empirically optimized values for the model parameters C (C = 30) and gamma (gamma = 5) as reported by Zhang et al. (2014) in their validation and optimization of the SVR-LSM method.

2.9. Exploratory deterministic tractography

Exploratory deterministic tractography was performed to characterize the WM pathways likely affected by lesions located near/in the WM that were found to strongly predict performance deficits in the SVR-LSM + DTLVC analyses. To identify these regions, up to three SVR-LSM peaks located in/near the WM (defined as voxels having a WM tissue probability of at least 20% as determined by the WM TPM included in SPM12) were defined as ROIs using the default ROI creation (12 mm spherical ROIs masked to exclude non-significant voxels) and peak reporting (top 3 peaks with minimum distances of 30 mm) settings in the *bspmview* toolbox for SPM12 (http://www.bobspunt.com/bspmview/). This resulted in up to three ROIs corresponding to the 3 peak WM predictors for each language outcome. We note that the WM TPM included in SPM12 (i.e. the mask used to identify the voxels used to define these ROIs) does not provide any information about the fiber tracts that pass through a given voxel, but rather quantifies the prior probability that a given voxel is WM. Thus, these ROIs were created solely on the basis of the results from the SVR-LSM analyses, and were intended to characterize the tracts that would be expected to be affected by damage to the regions identified by the SVR-LSM analyses. These SVR-LSM-derived ROIs should not be confused with the bottleneck ROIs described earlier, as they were defined independently of the bottleneck ROIs used for the a priori regression analyses. We further emphasize that like previous post-hoc tract analyses described in the Introduction (Mirman et al., 2015a; Turken and Dronkers, 2011), these analyses cannot be used to directly infer that damage to the identified tracts predicts task performance, and are provided only to characterize the potentially affected tracts so as to assess whether they are consistent with what is expected based on our *a priori* analyses and to inform future hypothesis-driven studies.

Deterministic fiber tracking utilized a publicly available group-averaged tractography atlas (WU-Minn HCP Consortium; HCP-842 atlas - http://dsi-studio.labsolver.org/download-images/hcp-842-template) from 842 healthy individuals' diffusion MRI data from the Human Connectome Project (2015 Q4, 900 subject release). Data were accessed under the WU-Minn HCP open access data use term. Data acquisition utilized a multi-shell diffusion scheme (b-values: 1000, 2000, and 3000 s/mm2; diffusion sampling directions: 90, 90, and 90; in-plane resolution: 1.25 mm). Data were reconstructed in MNI template space using Q-space diffeomorphic reconstruction (QSDR) as implemented in DSI_Studio (Yeh and Tseng, 2011) to obtain the spin distribution function (Yeh et al., 2010) (diffusion length sampling ratio: 1.25; output resolution: 2 mm). Deterministic fiber tracking (Yeh et al., 2013) proceeded by seeding the whole brain to calculate 100,000 tracts terminating within the 20% thresholded GM TPM included with SPM12, and used default tracking parameters implemented in DSI_studio (angular threshold: 60°; step size: 1 mm; quantitative anisotropy threshold determined automatically by DSI Studio to be 0.24; tracks with length <30 mm were discarded). The resulting tracts were filtered to leave only tracts that passed through each ROI. The filtered tracts were manually separated and labeled according to previous reports (Catani et al., 2002; Catani et al., 2013; Catani and Mesulam, 2008; Catani and Thiebaut de Schotten, 2008; Hua et al., 2008; Turken and Dronkers, 2011).

3. Results

3.1. Region of interest multiple regression analyses

3.1.1. Verbal fluency

Inspection of fitted residual and normal probability plots of the multiple regression for the verbal fluency measure indicated violations of homoscedasticity. To correct for this, a constant of 1 was added to verbal fluency scores and the resulting values were transformed using a square root transformation; inspection of fitted residuals and normal probability plots indicated that this transformation successfully corrected the violated assumptions. Notably, the untransformed and transformed verbal fluency measures were highly correlated ($r = 0.97$), so the results of the analyses using the transformed variable can be interpreted in a straightforward way.

The 4-predictor model significantly predicted verbal fluency scores ($R^2 = 0.59$, $F_{4,38} = 13.7$, $p < 0.001$, corrected). Similar results were obtained when the model was fit to the untransformed scores ($R^2 = 0.58$, $F_{4,38} = 13.1$, $p < 0.001$, corrected). Anterior lesion status, posterior lesion status, and the interaction term each uniquely predicted verbal fluency scores (see Fig. 3A–B). Because significant effects were revealed for the anterior bottleneck ROI, posterior bottleneck ROI, and interaction term, we next fit the lesion load control model using the tract and cortical lesion load information for both the anterior and posterior ROIs. The lesion load control model included lesion loads on the SLF/AF, ILF, IFOF, UF, and ATR as predictors to account for the effects of concomitant damage to the tracts used to define each bottleneck ROI, and lesion loads on both sets of cortical control ROIs as predictors (14 total predictors – Fig. 3A). The lesion load control model predicted verbal fluency scores ($R^2 = 0.64$, $F_{14,28} = 3.52$, $p = 0.002$). Next, to determine if the significant bottleneck ROI predictors from the 4-predictor model provided additional information about verbal fluency scores, anterior ROI lesion status, posterior ROI lesion status, and the interaction term were added to the lesion load control model, resulting in the full model (Fig. 3A). The full model explained significantly more variance than the lesion load control model ($R^2 = 0.84$, $\Delta R^2 = 0.20$, $F_{3,25} = 10.42$, $p < 0.001$, corrected), indicating that bottleneck lesion status provided unique information about verbal fluency scores even after accounting for concomitant cortical and tract damage. Plots illustrating the difference in model fit between the lesion load control model and the full model are shown in Fig. 3C.

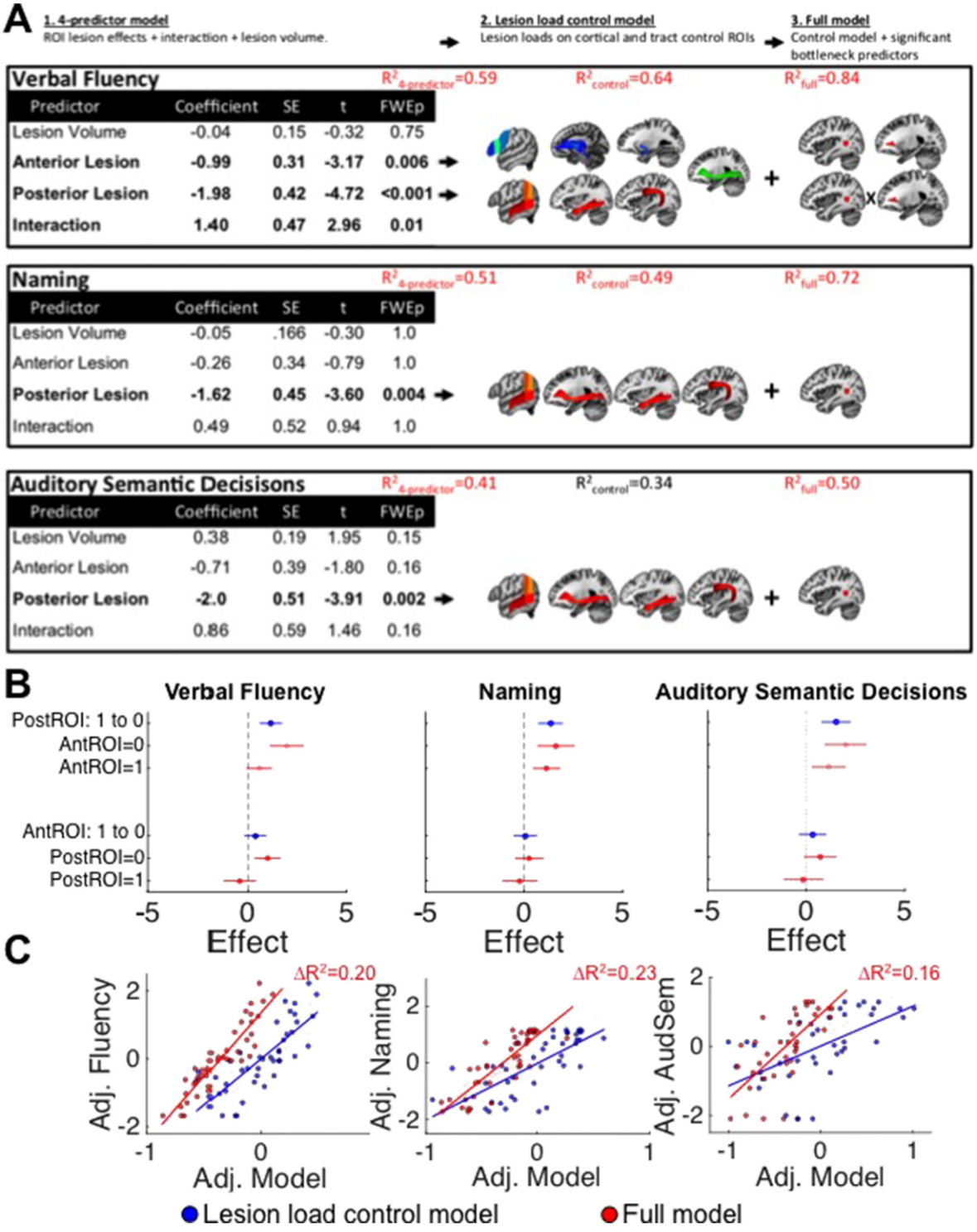

Fig. 3. Region of interest multiple regression results. **A**. For each language outcome, tables containing statistics for predictors from the 4-predictor model are shown (left – note: bold table text indicates $p < 0.05$, FWE-corrected), along with graphical depictions illustrating the variables entered in the first blocks (lesion load control model – middle) and second blocks (full model – right) of the hierarchical regressions performed to control for cortical and tract damage effects. The R^2 for each model is also shown (note: red R^2 text indicates $p < 0.05$, FWE-corrected). **B**. Bottleneck ROI effects from the 4-predictor models are shown. Blue data points illustrate the estimated effects (and 95% CIs) of changing lesion status at each bottleneck ROI from lesioned (1) to un-lesioned (0) on each language outcome, after averaging out the effects of other predictors. Red data points illustrate the estimated effects (and 95% CIs) for each bottleneck ROI when the other bottleneck ROI is lesioned (1) or unlesioned (0). Note: parameter estimates for each bottleneck ROI shown in (A) correspond to the effects of damage when the other ROI is not lesioned (i.e. =0). **C**. For each language outcome, regression fits for the lesion load control model (shown in blue) and full model (shown in red) are shown along with the change in R^2 between the two models (note: red R^2 text indicates $p < 0.05$, FWE-corrected).

3.1.2. Naming

The 4-predictor model significantly predicted naming scores ($R^2 = 0.51$, $F_{4,38} = 9.95$, $p < 0.001$, corrected). Only posterior lesion status uniquely predicted naming scores (see Fig. 3A–B). Because significant effects were revealed only for the posterior bottleneck ROI in the 4-predictor model, we next fit the lesion load control model using lesion loads for the tract and cortical controls for the posterior ROI. The lesion load control model included lesion loads on the SLF/AF, ILF, and IFOF as predictors to account for the effects of concomitant damage to the tracts used to define the posterior bottleneck ROI, and lesion loads on the posterior cortical control ROIs as predictors (8 total predictors – Fig. 3A). The lesion load control model predicted naming scores ($R^2 = 0.49$, $F_{8,34} = 4.04$, $p = 0.002$). Next, to determine if the significant posterior bottleneck ROI predictor from the 4-predictor model provided additional information about naming scores, posterior ROI lesion status was added to the lesion load control model, resulting in the full model (Fig. 3A). The full model explained significantly more variance than the lesion load control model ($R^2 = 0.72$, $\Delta R^2 = 0.23$, $F_{1,33} = 28.03$, $p < 0.001$, corrected), indicating that posterior ROI lesion status provided unique information about naming scores even after accounting for concomitant cortical and tract damage. Plots illustrating the difference in model fit between the lesion load control model and the full model are shown in Fig. 3C.

3.1.3. Auditory semantic decisions

The bottleneck 4-predictor model significantly predicted auditory semantic decision scores ($R^2 = 0.41$, $F_{4,34} = 5.91$, $p = 0.001$, corrected). Only posterior lesion status uniquely predicted auditory semantic decision scores (Fig. 3A–B). Because significant effects were revealed for the posterior bottleneck ROI in the bottleneck predictor model, we next fit the lesion load control model to lesion loads for the tract and cortical controls for the posterior ROI. The lesion load control model included lesion loads on the SLF/AF, ILF, and IFOF as predictors to account for the effects of concomitant damage to the tracts used to define the posterior bottleneck ROI, and lesion loads on the posterior cortical control ROIs as predictors (8 predictors – Fig. 3A). The control model did not significantly predict auditory semantic decision scores ($R^2 = 0.34$, $F_{8,30} = 1.91$, $p = 0.09$). Next, to determine whether the significant posterior bottleneck ROI predictor from the 4-predictor model provided additional information about auditory semantic decision scores, posterior ROI lesion status was added to the lesion load control model, resulting in the full model (Fig. 3A). The full model explained significantly more variance than the lesion load control model ($R^2 = 0.50$, $\Delta R^2 = 0.16$, $F_{1,29} = 9.28$, $p = 0.005$, corrected), indicating that ROI lesion status provided unique information about auditory semantic decision scores even after accounting for concomitant cortical and tract damage. Plots illustrating the difference in model fit between the lesion load control model and the full model are shown in Fig. 3C.

3.2. Multivariate lesion symptom mapping with lesion volume control

Significant lesion predictors of deficits in each of the language measures are shown in Fig. 4A, and cluster peaks for clusters with at least 10 voxels are provided in the Table. Fig. 4B illustrates the overlap of the SVR-LSM FDR maps with each of the bottleneck ROIs.

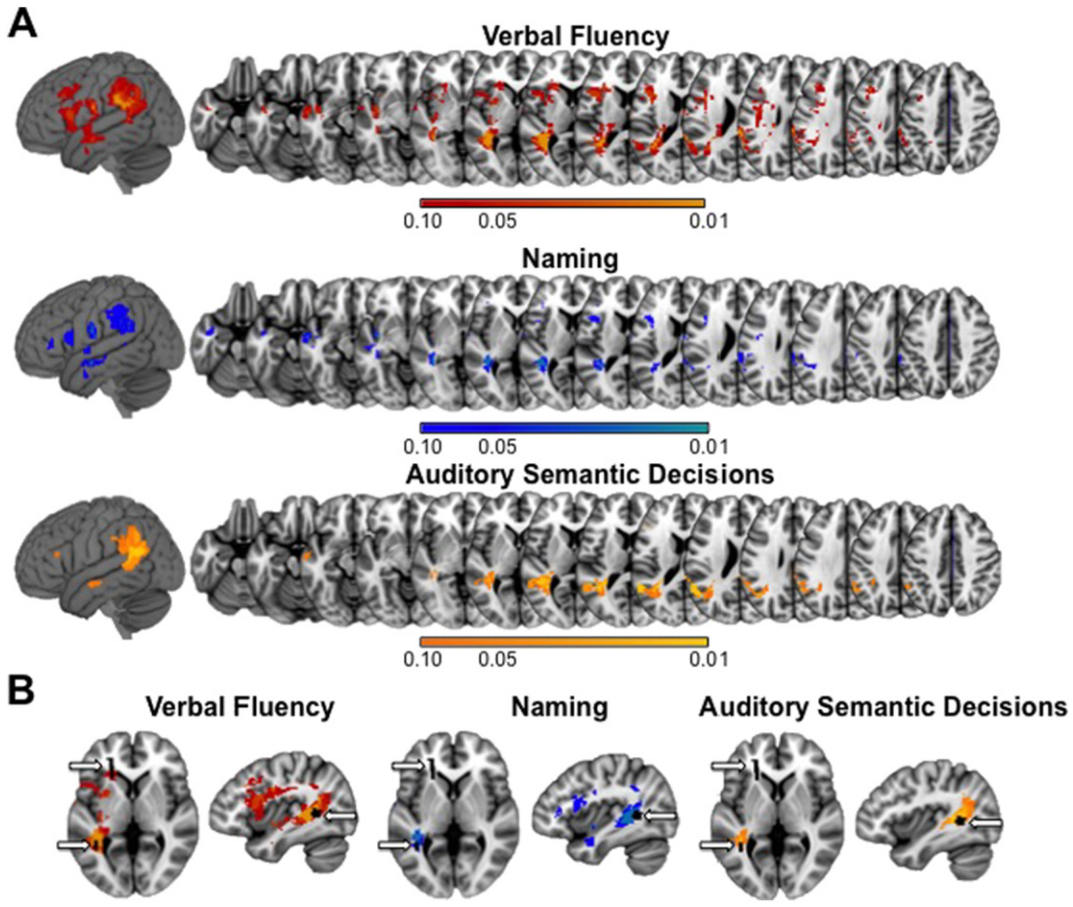

Fig. 4. Multivariate lesion-symptom-mapping results. **A**. Left hemisphere lesion predictors of deficits in verbal fluency (top), naming (middle), and auditory semantic decision (bottom) abilities. Each map shows FDR-corrected *p*-values (ranging from 0.1 to 0.01) obtained from permutation testing of the SVR-LSM + DTLVC model (2000 permutations). B. Overlap of left hemispheric lesion predictors of language deficits identified by SVR-LSM + DTLVC and white matter bottleneck ROIs. White matter bottleneck ROIs are shown in black and indicated by arrows.

Note that while both ROIs overlapped with the map for verbal fluency, only the posterior bottleneck ROI overlapped with the maps for naming and auditory semantic decisions, corroborating the results of our *a priori* analyses.

Three WM peaks meeting our criteria were identified for verbal fluency and naming, and one WM peak was identified for auditory semantic decisions. These were used for exploratory tractography analyses as described in the Methods.

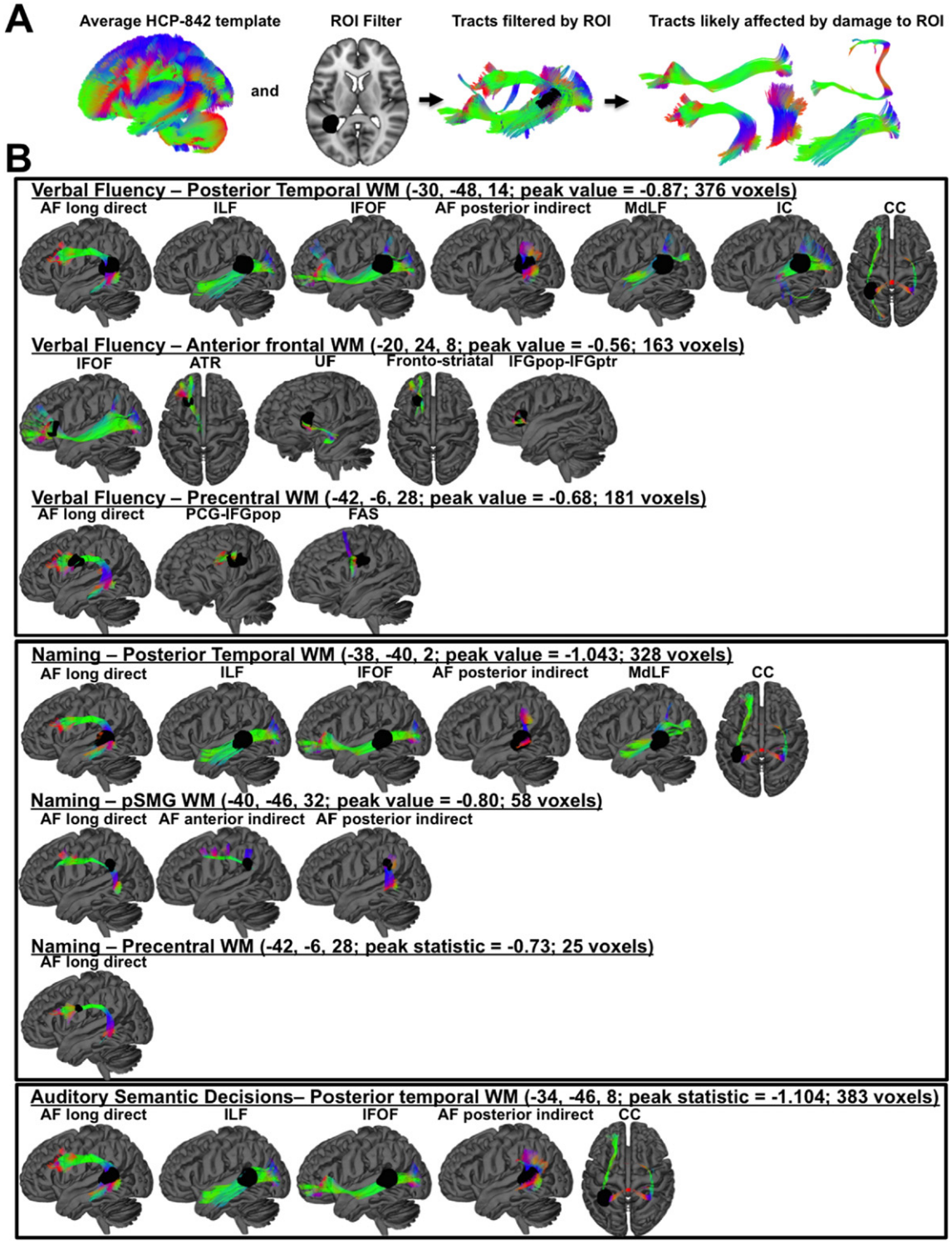

Fig. 5. Deterministic tractography results. **A**. A schematic diagram illustrating the deterministic tractography approach is shown. Fibers from the average HCP-842 template with termination points in the grey matter were filtered to leave only fibers that passed through each left hemisphere ROI. The resulting set of fibers was then segmented into constituent tracts that were labeled according to previous reports. **B**. 3D brain renderings show each peak white matter ROI (black regions) and the fiber tracts identified by the deterministic tractography analyses. Tracts associated with three ROIs are shown for Verbal Fluency (top panel) and Naming (middle panel), and tracts associated with one ROI are shown for Auditory Semantic Decisions (bottom panel). *Note: AF-Arcuate fasciculus; ILF-Inferior longitudinal fasciculus; IFOF-Inferior fronto-occipital fasciculus; MdLF-Middle longitudinal fasciculus; IC-Internal capsule; CC-Corpus callosum; ATR-Anterior thalamic radiations; UF-Uncinate fasciculus; IFGpop-Inferior frontal gyrus pars opercularis; IFGptr-Inferior frontal gyrus pars triangularis; PCG-Precentral gyrus; FAS-Frontal aslant tract*

3.3. Deterministic tractography

As described in the Methods section, ROIs were created on the white matter peak regions for each SVR-LSM + DTLVC analysis, and deterministic tractography was performed to identify tracts passing through each ROI in the HCP-842 healthy control template. The approach is outlined in Fig. 5A. Tracts identified for each ROI are shown in Fig. 5B along with the peak coordinates and statistics for the corresponding ROI from the SVR-LSM + DTLVC analysis.

4. Discussion

4.1. White matter bottleneck damage contributes to chronic language deficits

Previous studies indicate that white matter bottlenecks exist near regions where damage predicts chronic language deficits in left hemispheric stroke patients. However, we are not aware of any studies that have explicitly shown that damage to the bottlenecks in these regions is itself predictive of chronic language deficits after left hemispheric stroke. This limits the strengths of inferences that can be drawn regarding the contribution of damage affecting these regions to chronic language deficits after stroke. The ROI analyses employed here, by contrast, (1) directly assessed how chronic language deficits relate to the lesioning of white matter bottlenecks in these areas that were identified *a priori* based on these previous reports, and (2) demonstrate that damage to these regions predicts chronic language deficits beyond what can be attributed to the overall amount of damage sustained by the tracts that pass through them and by nearby language-relevant cortical areas. The results of our data-driven analyses using SVR-LSM + DTLVC were consistent with the results of our ROI analyses. Our data build upon previous reports that implicate white matter bottleneck lesions as potential contributors to chronic language deficits after left hemispheric stroke (Mirman et al., 2015a; Turken and Dronkers, 2011), and support the broader notion that regions of high tract overlap represent critical anatomical areas that may have a particularly negative impact on cognitive function if damaged (Corbetta et al., 2015).

Naming, verbal fluency, and auditory semantic decision scores were significantly impaired by damage to the posterior bottleneck region (Figs. 3 and 4). This corroborates findings from previous studies that have implicated damage to the WM underlying the posterior temporal lobe in language deficits that span broad domains of language function (Baldo et al., 2013; Bonilha et al., 2015; Butler et al., 2014; Dronkers et al., 2004; Fridriksson, 2010; Fridriksson et al., 2013; Geva et al., 2012; Harvey and Schnur, 2015; Henseler et al., 2014; Ivanova et al., 2016). This region contains fibers associated with both dorsal (AF) and ventral (ILF and IFOF) language processing streams that support integrated sensorimotor language processing and the extraction of meaning from auditory language, respectively (Kümmerer et al., 2013; Saur et al., 2008). Thus, our results support the notion that damage to the WM underlying the left posterior temporal lobe is associated with broad language impairments because it can simultaneously disrupt both dorsal and ventral language pathways. Notably, a recent diffusion-weighted imaging study of chronic left hemispheric stroke patients with aphasia reported that a patient with damage confined to the portion of the left posterior temporo-parietal WM containing fibers associated with these tracts (i.e. AF, IFOF, and ILF) showed profound impairments of speech production despite an absence of frontal damage, providing single-case evidence congruent with our conclusion (Ivanova et al., 2016). Our SVR-LSM + DTLVC analyses also found that damage to the WM in this region is a strong predictor of deficits in all three language outcomes (Fig. 4; Table 1). Congruent with the findings of Turken and Dronkers (2011), our *post-hoc* fiber tracking results suggest that damage to these posterior temporo-parietal WM regions may affect multiple tracts that include the AF long direct segment, AF posterior indirect segment, ILF, IFOF, and tapetum of the corpus callosum (Fig. 5B). While speculative, the incorporation of information regarding WM lesioning in posterior temporo-parietal regions could aid the development of future prognostic criteria for patients with left hemispheric stroke.

Table 1
SVR-LSM + DTLVC peak and cluster statistics.

Peak location	Extent	beta	x	y	z
Cluster peaks for SVR-LSM + DTLVC of verbal fluency deficits					
Posterior middle temporal WM	3869	−0.87	−30	−48	14
Precentral WM	3869	−0.68	−42	−6	28
Posterior supramarginal gyrus	3869	−0.58	−60	−46	26
Callosal WM	22	−0.48	−22	−12	28
Postcentral gyrus	179	−0.44	−64	−12	16
Anterior middle temporal gyrus	179	−0.37	−60	−8	−16
Callosal WM	15	−0.40	−22	−22	34
Precuneate WM	11	−0.31	−24	−48	42
Cluster peaks for SVR-LSM + DTLVC of naming deficits					
Posterior middle temporal WM	536	−1.04	−38	−40	2
Posterior supramarginal WM	536	−0.79	−40	−46	32
Precentral WM	25	−0.73	−42	−6	28
Precentral gyrus	175	−0.71	−40	6	26
Postcentral gyrus	38	−0.71	−64	−12	16
Planum polare	72	−0.69	−42	−16	−4
Posterior supramarginal gyrus	164	−0.66	−58	−48	22
Anterior supramarginal gyrus	164	−0.41	−64	−24	40
Parietal operculum	32	−0.66	−46	−30	24
Anterior middle temporal WM	123	−0.65	−44	−2	−24
Inferior frontal gyrus PTr WM	15	−0.62	−38	34	4
Anterior middle temporal gyrus	80	−0.55	−60	−8	−16
Anterior superior temporal gyrus	10	−0.45	−62	−6	−2
Cluster peaks for SVR-LSM + DTLVC of semantic decision deficits					
Posterior middle temporal WM	1475	−1.10	−34	−46	8
Posterior supramarginal gyrus	1475	−0.90	−58	−46	28
Inferior frontal gyrus PTr	11	−0.68	−52	26	18
Anterior middle temporal gyrus	33	−0.59	−60	−8	−16

While we found that damage to the anterior bottleneck predicted deficits in verbal fluency, we did not find a relationship between damage to this region and deficits in picture naming or auditory semantic processing as measured in this study. We note that Mirman et al. (2015a, 2015b) found that deficits in semantic recognition, a variable obtained from a factor analysis of scores on a large battery of neuropsychological language measures that featured the strongest loadings from verbal and non-verbal semantic category judgment tests and picture naming tests, were predicted by damage to a white matter cluster partially overlapping with the anterior bottleneck region. While we did not find evidence for a relationship between damage to this region and performance on the BNT (a picture naming test), we found evidence that damage to this region may contribute to verbal category fluency deficits (Figs. 3 and 4). Notably, the measures employed by Mirman et al. (2015a, 2015b) did not include measures analogous to the verbal fluency measures utilized in this study. One possibility is that damage to this region contributes to more general deficits in categorical processing, as this might result in deficits in both verbal fluency tasks using category prompts (i.e. as implemented here) and in tasks requiring semantic category judgments (i.e. as implemented in the study by Mirman and colleagues). The auditory semantic decision task used here also requires semantic category judgments, and while neither our ROI-driven analysis nor our SVR-LSM + DTLVC analysis detected significant relationships between auditory semantic decision performance and damage to this region at corrected thresholds, we note that the effect for the ROI-analysis showed weak near-significance at a per-comparison threshold (uncorrected $p = 0.08$), and partial overlap between this region and the SVR-LSM + DTLVC map for auditory semantic decisions (but not for naming) was also observed when no cluster-forming threshold (i.e. $k > 0$ applied to the FDR-corrected map) or a liberal voxel-wise threshold (i.e. $p < 0.05$, uncorrected) was employed (see Supplementary Fig. 4). The interpretation that damage to this region may impair categorical rather than strictly semantic processing is also supported by our

supplementary analyses showing that damage to this region predicts deficits in both phonological and semantic category verbal fluency (Supplementary Analyses 1 and 2). While this suggests that deficits following damage to this region are not confined to semantic processes, this cannot be unequivocally stated based on these data and this should be investigated by future studies.

Finally, one of the three peak WM lesion predictors of verbal fluency deficits identified by the SVR-LSM + DTLVC analyses partially overlapped with the anterior bottleneck region (Fig. 4B). The fiber tracking of the ROI derived from this region provides some further clues as to why damage to this region might contribute to verbal fluency deficits. In addition to fibers associated with the left ATR, UF, and IFOF (i.e. the tracts used to define the anterior bottleneck ROI), fibers associated with the left FAS, cortico-striatal projections, and short-range cortico-cortico fibers connecting left IFGpop and IFGptr also pass through this region (Fig. 5B). Previous research indicates that verbal fluency deficits in patients with primary progressive aphasia likely result in part from the degeneration of the FAS, which connects the superior frontal gyrus/pre-supplementary motor area to the IFGpop (Catani et al., 2013). Deficits in both letter and semantic fluency have been reported following damage to basal ganglia structures that include the caudate and putamen (Biesbroek et al., 2015; Chouiter et al., 2016), and basal ganglia degeneration correlates with verbal fluency deficits in individuals with HIV (Thames et al., 2012). FMRI evidence also implicates the basal ganglia in aspects of language processing that include word generation and speech fluency (Lu et al., 2010; Seghier et al., 2014; Seghier and Price, 2010). Furthermore, inhibitory signaling from the left IFGpop to left IFGptr likely occurs during word generation, and may reflect the application of phonological constraints to lexical retrieval processes during word generation (Heim et al., 2009). Thus, in addition to the ventral pathway and thalamo-cortical fibers passing through the anterior bottleneck region, fibers associated with at least three pathways relevant to verbal fluency may also traverse this region. While these conclusions should be considered tentative until confirmed by future studies, these results can inform hypotheses about the pathways that may contribute to verbal fluency deficits following damage to this region.

4.2. Additional lesion predictors of language deficits

Our SVR-LSM + DTLVC analyses revealed significant associations between damage to portions of the left pSMG/AG/IFG/ATL/pMTG and deficits in all three language outcomes, although the precise locations and extents differed for each outcome (Fig. 4). Owing to the coarser nature of our language measures relative to those utilized by other recent lesion-deficit studies (Mirman et al., 2015a; Mirman et al., 2015b), the ubiquity of these regions as lesion predictors of deficits in the current study may reflect involvement in processes "shared" by all three measures (e.g. associative retrieval, verbal working memory, etc.) (Lau et al., 2015). This may also, in part, reflect the role of some of these regions as heteromodal "information convergence zones" that are important for general semantic processing (Binder and Desai, 2011). Indeed, recent evidence suggests that the pMTG/AG/ATL act as cortical hubs for information convergence among control (i.e. fronto-parietal network), language (i.e. perisylvian network), and memory-oriented (i.e. default-mode network) modules of the larger semantic network, while the left pSMG and IFG act as cortical hubs for information convergence within language-oriented and control-oriented networks, respectively (Xu et al., 2016). Thus, these regions may act as cortical interfaces that enable within- and between-network information convergence during language processing. Damage to these regions could potentially disrupt the interactions among the network submodules that together form the larger-scale semantic network, and this could produce broad language impairments. However, this cannot be confirmed by these data, and should be considered speculative until confirmed by future studies.

Briefly, we note that the lesion predictors of naming deficits identified here are highly consistent with those of another recent study that investigated the lesion correlates of overall performance deficits on the same picture naming test we employed (Baldo et al., 2013). Baldo et al. (2013) found that deficits in picture naming were predicted by lesions affecting the anterior, middle, and posterior segments of the left MTG/STG and the underlying WM. Other recent studies have also implicated damage to the left anterior temporal lobe, IFG, and pSMG/AG in post-stroke naming deficits (Lau et al., 2015; Yourganov et al., 2016; Pustina et al., 2016), but damage to the left posterior MTG/STG and underlying white matter is consistently implicated in post-stroke naming deficits (Baldo et al., 2013; Henseler et al., 2014; Pustina et al., 2016; Yourganov et al., 2016). Recent lesion-mapping (Fridriksson, 2010) and structural connectivity (Bonilha et al., 2015) studies also indicate that lesions affecting the cortex and/or structural connections of these regions are associated with poorer responses to therapeutic interventions in patients with anomia. Speculatively, this may reflect the both the structural bottleneck properties of the WM underlying these regions and their role as cortical language network hubs (Turken and Dronkers, 2011; Xu et al., 2016).

The lesion predictors of verbal fluency deficits identified here are also largely consistent with the results of previous studies, as lesions to the IFG, middle frontal gyrus (MFG), PCG, anterior AF, anterior insula, and basal ganglia have previously been implicated in verbal fluency and speech fluency deficits after stroke (Bates et al., 2003; Biesbroek et al., 2015; Chouiter et al., 2016; Fridriksson et al., 2013; Yourganov et al., 2016). As noted in the Methods section, there is evidence for only a partial overlap in the neural systems supporting semantic and letter category fluency (Biesbroek et al., 2015; Chouiter et al., 2016; Costafreda et al., 2006; Heim et al., 2009), with previous lesion reports suggesting differences in temporal/parietal and frontal involvement, respectively (Baldo et al., 2006; Chouiter et al., 2016). Unfortunately, the high correlation between deficits on the semantic and letter fluency tasks observed in this sample made it ill-suited for the purpose of differentiating lesion predictors of the two tasks. Indeed, the results of our supplementary ROI (Supplementary Analysis 1) and SVR-LSM + DTLVC (Supplementary Analysis 2) analyses that were performed separately for each task were highly similar to those obtained for the general verbal fluency measure, with small differences in temporal/parietal vs. inferior frontal contributions to semantic and letter fluency, respectively (Supplementary Fig. 1). We therefore stress that our findings may primarily reflect regions that support processes common to both tasks.

Finally, while the auditory semantic decision task has been studied extensively in both healthy individuals and neurological patients using fMRI (Binder et al., 1997; Binder et al., 2008; Donnelly et al., 2011; Eaton et al., 2008; Frost et al., 1999; Griffis et al., 2016b; Kim et al., 2011; Szaflarski et al., 2002; Szaflarski et al., 2008; Szaflarski et al., 2011; Szaflarski et al., 2013), we are not aware of any other studies that have investigated the lesion correlates of performance on this task. Previous studies using different task paradigms have implicated damage to the left IFG, pSTG/pMTG, and pSMG/AG, in deficits of auditory comprehension and/or discrimination (Bates et al., 2003; Dronkers et al., 2004; Fridriksson et al., 2013; Geva et al., 2012; Pustina et al., 2016; Yourganov et al., 2016), and have implicated anterior temporal lobe lesions in semantic deficits when controlling for deficits in comprehension (Mirman et al., 2015a; Mirman et al., 2015b; Schwartz et al., 2009; Walker et al., 2011). Importantly, all of these regions were identified by our analysis, and are associated with the semantic network recruited by this task (Binder et al., 1997; Binder et al., 2009). Further, while cortical lesion predictors of deficits on this task were not found to be strongly localized to the superior temporal sulcus (BA22 – STS), previous reports have suggested that the posterior middle temporal cortex may be more important for auditory semantic comprehension than the canonical Wernicke's area (Dronkers et al., 2004).

4.3. Spatially specific white matter damage and controls for concomitant damage

The results from our ROI analyses indicate that the observed relationships between chronic language deficits and lesions affecting white matter bottlenecks are not attributable to the overall amount of damage sustained by the tracts that pass through them or by the overall amount of damage to nearby cortical language areas (Fig. 3). As previously noted, few studies investigating relationships between tract damage and language outcome have attempted to control for concomitant damage to other tracts or for cortical damage, weakening the conclusions that can be drawn regarding the effects of damage to the tracts being investigated. As our findings indicate that the effects of damage to long-range white matter pathways on language outcomes may show some spatial specificity, it is worth considering that spatially non-discriminative measures such as total tract lesion load may not be ideal for characterizing lesion-symptom relationships. Indeed, a similar argument regarding the importance of developing more informative characterizations of tract damage for lesion-behavior analyses was made by recent study comparing measures of lesion load and expected tract disconnection as predictors of chronic deficits in naming and verbal fluency (Hope et al., 2015). The importance of considering spatial specificity in analyses of tract damage is also supported by recently reported data from Ivanova et al. (2016). Their data suggest that the relationship between fractional anisotropy (FA) in portions of the AF and ILF near the posterior temporal lobe shows a stronger relationship to language deficits in chronic stroke patients than FA in other portions of these tracts, and also suggest that FA in the IFOF shows stronger relationships to language in portions slightly anterior to this location (Ivanova et al., 2016). In the current study, despite the considerable explanatory power of some of the lesion load control models (Fig. 3A), the spatially specific lesion status information for the relevant white matter bottleneck ROIs significantly improved all models as indicated by the R^2 change analysis (Fig. 3).

However, it is important to note that it remains possible that the bottleneck regions correspond to vital portions of only one or two tracts that pass through them. While determining whether spatially specific lesion-behavior relationships reflect the contribution of a single or multiple tracts is beyond the scope of the current study, this should be acknowledged as a possibility. Nonetheless, our study and other recent studies emphasize the importance of considering factors beyond overall lesion load in analyses of tract damage. Future studies that employ more spatially discriminative measures of tract damage, along with measures of tract disconnection, are important for coming to a better understanding of how post-stroke deficits relate to disruptions of inter-regional communication.

4.4. Limitations and conclusions

A limitation of the current study is that the bottleneck ROIs were defined using thresholded tract probability maps. Thus, it might be argued that the results of our ROI analyses may differ depending on threshold choice. However, the fact that similar results were obtained from a data-driven analysis (SVR-LSM) and *post-hoc* fiber tracking indicates that our results are robust to changes in analytical strategy. This is also supported by our supplemental analyses employing alternate ROI definitions (Supplementary Material 4). However, future studies using direct white matter measurements (i.e. using diffusion imaging data) are necessary to fully characterize how these regions contribute to post-stroke language deficits. Another limitation is that only chronic patients were included, and this prevents drawing strong conclusions about how damage to the bottleneck regions affects the success of long-term recovery vs. chronic aphasia severity, per se (i.e. it is unclear if damage to these regions causes severe long-term deficits, whether it impedes recovery, or both). Studies in patients presenting with acute symptoms and longitudinal studies of recovery are necessary allow for such conclusions. In addition, only 4 patients had damage to only the posterior bottleneck ROI, and so future studies that incorporate more patients with damage to this region are necessary to fully characterize its relationship to post-stroke language function. It is also important to note that, with the exception of the BNT, basic language task data (e.g. Western Aphasia Battery sub-tests) were not available for all patients. Indeed, two out of the three (i.e. verbal fluency and auditory semantic decisions) language measures utilized in this study are relatively complex language tasks that rely on high-level cognitive processes, and so inferences about the specific language sub-processes that are disrupted in patients with poor performance on these tasks are not possible. Future studies that incorporate more extensive language testing are a necessary next step to understand the specific language processes that are disrupted by damage to the regions studied here. Lastly, it is not clear how damage to these regions affects language network function, although it would be expected that associated disruptions of connectivity, particularly by damage to posterior temporo-parietal areas, might impede the recovery of typical function in left hemispheric language networks. This should be addressed by future studies.

Our data support the conclusion that WM bottlenecks correspond to structural vulnerabilities in the neural architecture of the distributed language network. Lesions affecting the posterior bottleneck correlated with chronic deficits on multiple language tasks. Lesions affecting the anterior bottleneck were primarily associated with chronic deficits in verbal fluency, although future work is needed to fully understand the specific effects of damage to this region. These results integrate with the results of other recent studies to emphasize the detrimental nature of damage to long-range white matter tracts on cognitive functions after stroke.

Funding

This work was supported by the National Institutes of Health [NIH R01 HD068488; NIH R01 NS048281].

Acknowledgements

Amber Martin.
Christi Banks.
Michel Thiebaut de Schotten for helpful discussion and for providing the atlas labels used for alternate ROI definitions in the supplementary analyses.
Two Anonymous Reviewers for their helpful criticisms and suggestions.
Data were provided [in part] 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.nicl.2017.02.019.

References

Aickin, M., Gensler, H., 1996. Adjusting for multiple testing when reporting research results: the Bonferroni vs Holm methods. Am. J. Public Health 86, 726–728.
Almairac, F., Herbet, G., Moritz-Gasser, S., de Champfleur, N.M., Duffau, H., 2014. The left inferior fronto-occipital fasciculus subserves language semantics: a multilevel lesion study. Brain Struct. Funct. 220, 1983–1995.
Alvarez, J.A., Emory, E., 2006. Executive function and the frontal lobes: a meta-analytic review. Neuropsychol. Rev. 16, 17–42.
Baldo, J.V., Arvalo, A., Patterson, J.P., Dronkers, N.F., 2013. Grey and white matter correlates of picture naming: evidence from a voxel-based lesion analysis of the Boston naming test. Cortex 49, 658–667.

Baldo, J.V., Schwartz, S., Wilkins, D., Dronkers, N.F., 2006. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. J. Int. Neuropsychol. Soc. 12, 896–900.

Bates, E., Wilson, S.M., Saygin, A.P., Dick, F., Sereno, M.I., Knight, R.T., Dronkers, N.F., 2003. Voxel-based lesion-symptom mapping. Nat. Neurosci. 6, 448–450.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B 57, 289–300.

Berker, E.A., Berker, A.H., Smith, A., 1986. Translation of Broca's 1865 report. Localization of speech in the third left frontal convolution. Archives of Neurology.

Biesbroek, J.M., van Zandvoort, M.J.E., Kappelle, L.J., Velthuis, B.K., Biessels, G.J., Postma, A., 2015. Shared and distinct anatomical correlates of semantic and phonemic fluency revealed by lesion-symptom mapping in patients with ischemic stroke. Brain Struct Funct.

Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M., Prieto, T., 1997. Human brain language areas identified by functional magnetic resonance imaging. J. Neurosci. 17, 353–362.

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cogn. Sci. 15, 527–536.

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. Cereb. Cortex 19, 2767–2796.

Binder, J.R., Swanson, S.J., Hammeke, T.A., Sabsevitz, D.S., 2008. A comparison of five fMRI protocols for mapping speech comprehension systems. Epilepsia 49, 1980–1997.

Bizzi, A., Nava, S., Ferrè, F., Castelli, G., Aquino, D., Ciaraffa, F., Broggi, G., DiMeco, F., Piacentini, S., 2012. Aphasia induced by gliomas growing in the ventrolateral frontal region: assessment with diffusion MR tractography, functional MR imaging and neuropsychology. Cortex 48, 255–272.

Blasi, V., Young, A.C., Tansy, A.P., Petersen, S.E., Snyder, A.Z., Corbetta, M., 2002. Word retrieval learning modulates right frontal cortex in patients with left frontal damage. Neuron 36, 159–170.

Bonilha, L., Gleichgerrcht, E., Nesland, T., Rorden, C., Fridriksson, J., 2015. Success of anomia treatment in aphasia is associated with preserved architecture of global and left temporal lobe structural networks. Neurorehabil Neural Repair.

Bose, A., Wood, R., Kiran, S., 2016. Semantic fluency in aphasia: clustering and switching in the course of 1 minute. Int. J. Lang. Commun. Disord. 0, 1–12.

Butler, R.A., Lambon Ralph, M.A., Woollams, A.M., 2014. Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. Brain 3248–3266.

Catani, M., Howard, R.J., Pajevic, S., Jones, D.K., 2002. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. NeuroImage 17, 77–94.

Catani, M., Mesulam, M., 2008. The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. Cortex 44, 953–961.

Catani, M., Mesulam, M.M., Jakobsen, E., Malik, F., Martersteck, A., Wieneke, C., Thompson, C.K., Thiebaut De Schotten, M., Dell'Acqua, F., Weintraub, S., Rogalski, E., 2013. A novel frontal pathway underlies verbal fluency in primary progressive aphasia. Brain 136, 2619–2628.

Catani, M., Thiebaut de Schotten, M., 2008. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 44, 1105–1132.

Chouiter, L., Holmberg, J., Manuel, A.L., Colombo, F., Clarke, S., Annoni, J.-M., Spierer, L., 2016. Partly segregated cortico-subcortical pathways support phonologic and semantic verbal fluency: a lesion study. Neuroscience 329, 275–283.

Corbetta, M., Ramsey, L., Callejas, A., Baldassarre, A., Hacker, C.D., Siegel, J.S., Astafiev, S.V., Rengachary, J., Zinn, K., Lang, C.E., Connor, L.T., Fucetola, R., Strube, M., Carter, A.R., Shulman, G.L., 2015. Common behavioral clusters and subcortical anatomy in stroke. Neuron 85, 927–941.

Costafreda, S.G., Fu, C.H.Y., Lee, L., Everitt, B., Brammer, M.J., David, A.S., 2006. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. Hum. Brain Mapp. 27, 799–810.

Donnelly, K.M., Allendorfer, J.B., Szaflarski, J.P., 2011. Right hemispheric participation in semantic decision improves performance. Brain Res. 1419, 105–116.

Dosenbach, N.U., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A.T., Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., Schlaggar, B.L., Petersen, S.E., 2007. Distinct brain networks for adaptive and stable task control in humans. Proc. Natl. Acad. Sci. U. S. A. 104, 11073–11078.

Dronkers, N.F., 1996. A new brain region for coordinating speech articulation. Nature 384, 159–161.

Dronkers, N.F., Wilkins, D.P., Van Valin, R.D., Redfern, B.B., Jaeger, J.J., 2004. Lesion analysis of the brain areas involved in language comprehension. Cognition 92, 145–177.

Eaton, K.P., Szaflarski, J.P., Altaye, M., Ball, A.L., Kissela, B.M., Banks, C., Holland, S.K., 2008. Reliability of fMRI for studies of language in post-stroke aphasia subjects. NeuroImage 41, 311–322.

Fama, M.E., Hayward, W., Snider, S.F., Friedman, R.B., Turkeltaub, P.E., 2017. Subjective experience of inner speech in aphasia: preliminary behavioral relationships and neural correlates. Brain Lang. 164, 32–42.

Fischl, B., Van Der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb. Cortex 14, 11–22.

Forkel, S.J., Thiebaut de Schotten, M., Dell'Acqua, F., Kalra, L., Murphy, D.G.M., Williams, S.C.R., Catani, M., 2014. Anatomical predictors of aphasia recovery: a tractography study of bilateral perisylvian language networks. Brain 137, 2027–2039.

Fridriksson, J., 2010. Preservation and modulation of specific left hemisphere regions is vital for treated recovery from anomia in stroke. J. Neurosci. 30, 11558–11564.

Fridriksson, J., Guo, D., Fillmore, P., Holland, A., Rorden, C., 2013. Damage to the anterior arcuate fasciculus predicts non-fluent speech production in aphasia. Brain 136, 3451–3460.

Friederici, A.D., Gierhan, S.M.E., 2013. The language network. Curr. Opin. Neurobiol. 23, 250–254.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

Frost, J.A., Binder, J.R., Springer, J.A., Hammeke, T.A., 1999. Language processing is strongly left lateralized in both sexes: evidence from MRI. Brain 122, 199–208.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. NeuroImage 15, 870–878.

Geranmayeh, F., Wise, R.J.S., Mehta, A., Leech, R., 2014. Overlapping networks engaged during spoken language production and its cognitive control. J. Neurosci. 34, 8728–8740.

Geva, S., Baron, J.-C., Jones, P.S., Price, C.J., Warburton, E.A., 2012. A comparison of VLSM and VBM in a cohort of patients with post-stroke aphasia. NeuroImage Clin. 1, 37–47.

Geva, S., Correia, M.M., Warburton, E.A., 2015. Contributions of bilateral white matter to chronic aphasia symptoms as assessed by diffusion tensor MRI. Brain Lang. 150, 117–128.

Griffis, J.C., Allendorfer, J.B., Szaflarski, J.P., 2016a. Voxel-based Gaussian naïve Bayes classification of ischemic stroke lesions in individual T1-weighted MRI scans. J. Neurosci. Methods 257, 97–108.

Griffis, J.C., Nenert, R., Allendorfer, J.B., Vannest, J., Holland, S., Dietz, A., Szaflarski, J.P., 2016b. The canonical semantic network supports residual language function in chronic post-stroke aphasia. Hum. Brain Mapp.

Harvey, D.Y., Schnur, T.T., 2015. Distinct loci of lexical and semantic access deficits in aphasia: evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. Cortex 67, 37–58.

Heim, S., Eickhoff, S.B., Amunts, K., 2009. Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. NeuroImage 48, 616–624.

Henseler, I., Regenbrecht, F., Obrig, H., 2014. Lesion correlates of patholinguistic profiles in chronic aphasia: comparisons of syndrome-, modality-and symptom-level assessment. Brain 137, 918–930.

Holland, R., Leff, A.P., Josephs, O., Galea, J.M., Desikan, M., Price, C.J., Rothwell, J.C., Crinion, J., 2011. Speech facilitation by left inferior frontal cortex stimulation. Curr. Biol. 21, 1403–1407.

Hope, T.M.H., Seghier, M.L., Prejawa, S., Leff, A.P., Price, C.J., 2015. Distinguishing the effect of lesion load from tract disconnection in the arcuate and uncinate fasciculi. Neuroimage 125:1169–1173. http://dx.doi.org/10.1016/j.neuroimage.2015.09.025.

Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C.M., Mori, S., 2008. Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. NeuroImage 39, 336–347.

Ivanova, M.V., Isaev, D.Y., Dragoy, O.V., Akinina Yu, S., Petryshevskii, A.G., Fedina, O.N., Shklovsky, V.M., Dronkers, N.F., 2016. Diffusion-tensor imaging of major white matter tracts and their role in language processing in aphasia. Cortex 1–17.

Jurado, M.B., Rosselli, M., 2007. The elusive nature of executive functions: a review of our current understanding. Neuropsychol. Rev. 17, 213–233.

Kaplan, E., Goodglass, H., Weintraub, S., 2001. The Boston Naming Test. Lippincott, Williams and Wilkins, Philadelphia, PA.

Kim, K.K., Karunanayaka, P., Privitera, M.D., Holland, S.K., Szaflarski, J.P., 2011. Semantic association investigated with functional MRI and independent component analysis. Epilepsy Behav. 20, 613–622.

Kozora, E., Cullum, C.M., 1995. Generative naming in normal aging: total output and qualitative changes using phonemic and semantic constraints. Clin. Neuropsychol. 9, 313–320.

Kümmerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Klöppel, S., Suchan, J., Karnath, H.O., Weiller, C., Saur, D., 2013. Damage to ventral and dorsal language pathways in acute aphasia. Brain 136, 619–629.

Lau, J.K.L., Humphreys, G.W., Douis, H., Balani, A., Bickerton, W., Rotshtein, P., 2015. The relation of object naming and other visual speech production tasks: a large scale voxel-based morphometric study. NeuroImage Clin 7, 463–475.

Lezak, M.D., Howieson, D.B., Loring, D.W., Hannay, J.H., Fischer, J.S., 1995. Neuropsychological Assessment. 3. Oxford University Press, New York.

Lu, C., Peng, D., Chen, C., Ning, N., Ding, G., Li, K., Yang, Y., Lin, C., 2010. Altered effective connectivity and anomalous anatomy in the basal ganglia-thalamocortical circuit of stuttering speakers. Cortex 46, 49–67.

Marchina, S., Zhu, L.L., Norton, A., Zipse, L., Wan, C.Y., Schlaug, G., 2011. Impairment of speech production predicted by lesion load of the left arcuate fasciculus. Stroke 42, 2251–2256.

Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O.K., Coslett, H.B., Schwartz, M.F., 2015a. Neural organization of spoken language revealed by lesion–symptom mapping. Nat. Commun. 6, 6762.

Mirman, D., Zhang, Y., Wang, Z., Coslett, H.B., Schwartz, M.F., 2015b. The ins and outs of meaning: behavioral and neuroanatomicsal dissociation of semantically-driven word retrieval and multimodal semantic recognition in aphasia. Neuropsychologia 76, 208–219.

Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A.W., Pike, G.B., Rosa, P., Evans, A., Zhang, J., Huang, H., Miller, M.I., Van PCM, Mazziotta J., Mori, S., Brain, F., 2008. Human brain white matter atlas: identification and assignment of common anatomical structures in superficial white matter. NeuroImage 43, 447–457.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113.

Peck, K.K., Moore, A.B., Crosson, B.A., Gaiefsky, M., Gopinath, K.S., White, K., Briggs, R.W., 2004. Functional magnetic resonance imaging before and after aphasia therapy: shifts in hemodynamic time to peak during an overt language task. Stroke 35, 554–559.

Perani, D., Cappa, S.F., Tettamanti, M., Rosa, M., Scifo, P., Miozzo, A., Basso, A., Fazio, F., 2003. A fMRI study of word retrieval in aphasia. Brain Lang. 85, 357–368.

Price, C.J., 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann. N. Y. Acad. Sci. 1191, 62–88.
Price, C.J., 2012. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. NeuroImage 62, 816–847.
Pustina, D., Coslett, H.B., Turkeltaub, P.E., Tustison, N., Schwartz, M.F., Avants, B., 2016. Automated segmentation of chronic stroke lesions using LINDA: lesion identification with neighborhood data analysis. Hum. Brain Mapp. 37:1405–1421. http://dx.doi.org/10.1002/hbm.23110.
Riecker, A., Mathiak, K., Wildgruber, D., Erb, M., Hertrich, I., Grodd, W., Ackermann, H., 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. Neurology 64, 700–706.
Rojkova, K., Volle, E., Urbanski, M., Humbert, F., Dell' Acqua, F., Thiebaut de Schotten, M., 2016. Atlasing the frontal lobe connections and their variability due to age and education: a spherical deconvolution tractography study. Brain Struct. Funct. 221: 1751–1766. http://dx.doi.org/10.1007/s00429-015-1001-3.
Saur, D., Kreher, B.W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., Weiller, C., 2008. Ventral and dorsal pathways for language. Proc. Natl. Acad. Sci. U. S. A. 105, 18035–18040.
Schwartz, M.F., Kimberg, D.Y., Walker, G.M., Faseyitan, O., Brecher, A., Dell, G.S., Coslett, H.B., 2009. Anterior temporal involvement in semantic word retrieval: voxel-based lesion-symptom mapping evidence from aphasia. Brain 132, 3411–3427.
Seghier, M.L., Price, C.J., 2010. Reading aloud boosts connectivity through the putamen. Cereb. Cortex 20, 570–582.
Seghier, M.L., Bagdasaryan, J., Jung, D.E., Price, C.J., 2014. The importance of premotor cortex for supporting speech production after left capsular-putaminal damage. J. Neurosci. 34, 14338–14348.
Skipper-Kallal, L.M., Mirman, D., Olson, I.R., 2015. Converging evidence from fMRI and aphasia that the left temporoparietal cortex has an essential role in representing abstract semantic knowledge. Cortex 69, 104–120.
Swick, D., Ashley, V., Turken, A.U., 2008. Left inferior frontal gyrus is critical for response inhibition. BMC Neurosci. 9, 102.
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.
Szaflarski, J.P., Allendorfer, J.B., Banks, C., Vannest, J., Holland, S.K., 2013. Recovered vs. not-recovered from post-stroke aphasia: the contributions from the dominant and non-dominant hemispheres. Restor. Neurol. Neurosci. 31, 347–360.
Szaflarski, J.P., Holland, S.K., Jacola, L.M., Lindsell, C., Privitera, M.D., Szaflarski, M., 2008. Comprehensive presurgical functional MRI language evaluation in adult patients with epilepsy. Epilepsy Behav. 12, 74–83.
Szaflarski, J.P., Vannest, J., Wu, S.W., DiFrancesco, M.W., Banks, C., Gilbert, D.L., 2011. Excitatory repetitive transcranial magnetic stimulation induces improvements in chronic post-stroke aphasia. Med. Sci. Monit. 17, CR132–CR139.
Thames, A.D., Foley, J.M., Wright, M.J., Panos, S.E., Ettenhofer, M., Ramezani, A., Streiff, V., El-Saden, S., Goodwin, S., Bookheimer, S.Y., Hinkin, C.H., 2012. Basal ganglia structures differentially contribute to verbal fluency: evidence from human immunodeficiency virus (HIV)-infected adults. Neuropsychologia 50, 390–395.
Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., Farah, M.J., 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc. Natl. Acad. Sci. U. S. A. 94, 14792–14797.
Turken, A.U., Dronkers, N.F., 2011. The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. Front. Syst. Neurosci. 5, 1.
Vigneau, M., Beaucousin, V., Hervé, P.Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., Tzourio-Mazoyer, N., 2006. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. NeuroImage 30, 1414–1432.
Vitali, P., Abutalebi, J., Tettamanti, M., Rowe, J., Scifo, P., Fazio, F., Cappa, S.F., Perani, D., 2005. Generating animal and tool names: an fMRI study of effective connectivity. Brain Lang. 93, 32–45.
Walker, G.M., Schwartz, M.F., Kimberg, D.Y., Faseyitan, O., Brecher, A., Dell, G.S., Coslett, H.B., 2011. Support for anterior temporal involvement in semantic error production in aphasia: new evidence from VLSM. Brain Lang. 117, 110–122.
Whiteside, D.M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M.R., Roper, B., 2015. Verbal fluency: language or executive function measure? Appl. Neurophysiol. 9095, 1–6.
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.
Yeh, F.C., Wedeen, V.J., Tseng, W.-Y.I., 2010. Generalized q-sampling imaging. IEEE Trans. Med. Imaging 29, 1626–1635.
Yeh, F.C., Tseng, W.Y.I., 2011. NTU-90: a high angular resolution brain atlas constructed by q-space diffeomorphic reconstruction. NeuroImage 58, 91–99.
Yeh, F.C., Verstynen, T.D., Wang, Y., Fernández-Miranda, J.C., Tseng, W.Y.I., 2013. Deterministic diffusion fiber tracking improved by quantitative anisotropy. PLoS One 8, 1–16.
Yourganov, G., Fridriksson, J., Rorden, C., Gleichgerrcht, E., Bonilha, L., 2016. Multivariate connectome-based symptom mapping in post-stroke patients: networks supporting language and speech. J. Neurosci. 36, 6668–6679.
Zhang, Y., Kimberg, D.Y., Coslett, H.B., Schwartz, M.F., Wang, Z., Yongsheng, Z., Kimberg, D.Y., Coslett, H.B., Schwartz, M.F., Wang, Z., 2014. Multivariate lesion-symptom mapping using support vector regression. Hum. Brain Mapp. 35, 997.