



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Lesion correlates of auditory sentence comprehension deficits in post-stroke aphasia

**Citation for published version:**

Adezati, E, Thye, M, Edmondson-Stait, A, Szaflarski, JP & Mirman, D 2022, 'Lesion correlates of auditory sentence comprehension deficits in post-stroke aphasia', *Neuroimage: Reports*, vol. 2, no. 1, 100076. <https://doi.org/10.1016/j.ynirp.2021.100076>

**Digital Object Identifier (DOI):**

[10.1016/j.ynirp.2021.100076](https://doi.org/10.1016/j.ynirp.2021.100076)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Neuroimage: Reports

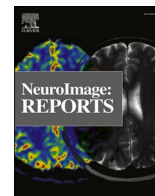
**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





## Lesion correlates of auditory sentence comprehension deficits in post-stroke aphasia

Erica Adezati<sup>a</sup>, Melissa Thye<sup>a</sup>, Amelia J. Edmondson-Stait<sup>b</sup>, Jerzy P. Szaflarski<sup>c</sup>, Daniel Mirman<sup>a,\*</sup>

<sup>a</sup> Department of Psychology, University of Edinburgh, Edinburgh, UK

<sup>b</sup> Centre for Clinical Brain Sciences, University of Edinburgh, Edinburgh, UK

<sup>c</sup> Department of Neurology and the University of Alabama at Birmingham Epilepsy Center, University of Alabama at Birmingham Heersink School of Medicine, Birmingham, AL, USA

### ARTICLE INFO

#### Keywords:

Aphasia  
Language comprehension  
Lesion-symptom mapping

### ABSTRACT

Auditory sentence comprehension requires coordination of multiple levels of processing: auditory-phonological perception, lexical-semantic comprehension, syntactic parsing and discourse construction, as well as executive functions such as verbal working memory (WM) and cognitive control. This study examined the lesion correlates of sentence comprehension deficits in post-stroke aphasia, building on prior work on this topic by using a different and clinically-relevant measure of sentence comprehension (the Token Test) and multivariate (SCCAN) and connectome-based lesion-symptom mapping methods. The key findings were that lesions in the posterior superior temporal lobe and inferior frontal gyrus (pars triangularis) were associated with sentence comprehension deficits, which was observed in both mass univariate and multivariate lesion-symptom mapping. Graph theoretic measures of connectome disruption were not statistically significantly associated with sentence comprehension deficits after accounting for overall lesion size.

### 1. Introduction

Auditory sentence comprehension requires the coordination of multiple levels of processing: auditory-phonological perception, lexical-semantic comprehension, syntax parsing and discourse construction, as well as executive functions such as verbal working memory (WM) and cognitive control. The phonological and lexical aspects appear to be supported by a “ventral stream” of regions in the lateral temporal lobe (e.g., Hickok and Poeppel, 2007), but the critical regions are less clear at the sentence level. Left inferior frontal cortex, inferior parietal cortex, and posterior superior temporal cortex are the most consistently implicated regions for sentence-level comprehension deficits (for a review see Wilson, 2017).

Early voxel-based lesion-symptom mapping (VLSM) studies used composite comprehension scores and found that posterior middle temporal gyrus (pMTG) damage was most strongly associated with comprehension deficits (Bates et al., 2003; Dronkers et al., 2004), possibly extending dorsally into angular gyrus (AG) and anteriorly to the anterior temporal lobe (ATL) and inferior frontal gyrus (IFG).

Subsequent VLSM studies have confirmed that pMTG and AG damage are associated with sentence comprehension deficits (Kristinsson et al., 2020; Pillay et al., 2017). Some of these studies have focused on so-called agrammatic sentence comprehension, which is characterised by substantially worse performance on unusual (“noncanonical”) sentence structures compared to more typical and familiar (“canonical”) sentence structures and is traditionally associated with Broca’s area damage (e.g., Grodzinsky and Santi, 2008). However, both canonical and noncanonical sentence comprehension appear to be associated with temporo-parietal damage (Rogalsky et al., 2018; Thothathiri et al., 2012). Lwi et al. (2021) compared three auditory comprehension tasks and found that pMTG damage was associated with deficits in all three tasks, along with differences between tasks: the lesion correlates of single word comprehension extended into inferior temporal regions and the lesion correlates of sequential commands extended into superior temporal and angular gyri.

Rogalsky et al. (2018) noted that damage to inferior frontal cortex was significantly associated with response bias in sentence plausibility judgments, suggesting that IFG may be involved in cognitive demands

\* Corresponding author.

E-mail address: [dan@danmirman.org](mailto:dan@danmirman.org) (D. Mirman).

<https://doi.org/10.1016/j.ynirp.2021.100076>

Received 9 October 2021; Received in revised form 27 December 2021; Accepted 28 December 2021

Available online 4 January 2022

2666-9560/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

related to sentence comprehension. One important factor is working memory (WM; which these studies seem to use synonymously with short-term memory [STM], although the two terms are not exactly equivalent in cognitive psychology). WM/STM is important for sentence comprehension because sentence input is a sequence of words that must be temporarily maintained in memory while constructing a holistic – usually hierarchical – representation of the meaning. Several studies have found that WM/STM deficits are associated with sentence comprehension deficits, and that both are associated with parietal-frontal damage (Barbey et al., 2014; Leff et al., 2009; Newhart et al., 2012).

The present study provides a useful converging perspective on sentence comprehension deficits. First, sentence comprehension was measured using the Token Test (De Renzi and Vignolo, 1962), which, to our knowledge, has not been used in prior lesion-symptom mapping studies. The sentences on the Token Test do not use non-canonical structures or complex embeddings, which allows an investigation of sentence comprehension decoupled from complex syntactic processing. The Token Test sentences also use a relatively restricted set of highly familiar concepts (colours, shapes, spatial and temporal relations), so deficits are unlikely to arise from impaired comprehension of individual words (i.e., lexical-semantic deficits).

However, the Token Test does require combinatorial processing in the form of maintaining and executing a sequence of movements. These maintenance and sequencing operations should rely on verbal WM. They may also require semantic control because repetition of items from semantic categories tends to produce competition that requires control resources to resolve, as seen in blocked cyclic naming and in “access” deficits more generally (Mirman and Britt, 2014). The Token Test is also used in clinical contexts to assess auditory sentence comprehension, so this study may provide helpful information for interpreting its results in clinical settings.

Second, there is growing concern about reproducibility, particularly regarding brain-behaviour relationships (e.g., Boekel et al., 2015), but direct replication of lesion-symptom studies is generally both practically and financially impossible (for more discussion see Geller et al., 2019). By using a different measure of sentence comprehension the present study provides a conceptual replication of prior LSM studies of sentence comprehension (particularly of Lwi et al., 2021, and Pillay et al., 2017). It further does this using fully reproducible methods and a publicly available data set (<https://osf.io/br3dm>).

Third, since sentence comprehension requires coordinating multiple processes, it may well rely on a network of non-contiguous regions. Standard mass-univariate VLSM methods are not very well-suited to detecting such networks. Thus, we supplement those analyses with a more recently developed multivariate LSM method based on sparse canonical correlations (SCCAN; Pustina et al., 2018), which is better able to detect when damage to multiple regions contribute to a deficit (for recent applications and additional discussion of the advantages of SCCAN see Thye et al., 2021; Thye and Mirman, 2018). To characterise the lesion impact on information transfer within networks, we further include graph-theoretical analyses of whether sentence comprehension deficits are associated with white matter connectivity disruption.

## 2. Methods

### 2.1. Data

The data were drawn from two large-scale studies of language processing following left hemisphere stroke. Detailed descriptions of the study design, participants, and neuroimaging protocols have been provided in previous studies (Allendorfer et al., 2012, 2021; Griffis et al., 2017; Nenert et al., 2017; Szaflarski et al., 2021). MRI data were collected using Philips 3T or Siemens 3T scanners. On Philips 3T scanner, the high-resolution T1-weighted anatomical scans were acquired using a magnetization prepared - rapid gradient echo (MPRAGE)

acquisition with the following parameters: TR/TE = 8100/3.7 ms, FOV = 25.0 × 21.0 × 18.0 cm, matrix = 252 × 211, flip angle = 8°, slice thickness = 1 mm. On Siemens Allegra 3T scanner, the corresponding scan had the following parameters: TR/TE = 2300/2.17 ms, FOV = 25.6 × 25.6 × 19.2 cm, matrix = 256 × 256, flip angle = 9°, slice thickness = 1 mm. Finally, the anatomical scan from Siemens Prisma 3T scanner used the following sequence: TR/TE = 2300/3.37 ms, FOV = 25.6 × 25.6 × 19.2 cm, matrix = 256 × 256, flip angle = 9°, slice thickness = 1 mm. The diffusion weighted imaging sequence parameters are provided in Table 2. Almost all of the participants included in the current study were previously included in a lesion-symptom mapping study of semantic and phonological fluency (Thye et al., 2021).<sup>1</sup>

Prospectively collected Magnetic Resonance Imaging (MRI) and psycholinguistic data from 50 participants with aphasia secondary to a single left hemisphere stroke were analysed. Analysis of these de-identified data was approved by the Institutional Review Board at the University of Alabama at Birmingham (IRB-120726004, IRB-120726006) and the PPLS Research Ethics panel at the University of Edinburgh (Ref No. 16–2021/2). Participant demographic information is presented in Table 1.

### 2.2. Auditory sentence comprehension measure

The behavioural measure for the present analyses was the Token Test (De Renzi and Vignolo, 1962), which requires participants to manually carry out the examiner’s verbal commands involving a set of tokens differing in shape, size and colour. The verbal instructions include simple commands (e.g., ‘Pick up the yellow rectangle’) as well as more complex ones (e.g., ‘Before touching the yellow circle, pick up the red rectangle’). Scores between 37 and 40 are thought to indicate a mild language impairment, scores ranging between 17 and 36 a moderate deficit, and scores less than 17 indicate severe comprehension deficits (as in previous work: Szaflarski et al., 2015). As shown in Table 1, the participants in this sample had a very broad range of time since stroke (2 months - 14 years); however, time since stroke was uncorrelated with Token Test performance ( $r = -0.065$ ,  $p > 0.65$ ), so it was not included in the statistical analyses.

**Table 1**  
Participant Demographic Information.

	Mean (SD)	Range
Age (years)	52.14 (15.12)	22.65–84.83
Lesion Size (cc)	90.97 (64.65)	1.88–262.83
Time Since Stroke (months)	38.87 (38.09)	2.24–167.93
Token Test	24.36 (13.22)	4–44
PPVT	189.35 (28.23)	75–223
BNT	33.88 (20.25)	0–60
SFT	19.12 (15.90)	0–62
COWAT	8.44 (8.64)	0–36
Gender (M:F)	33:17	
Handedness (R:L:A)	44:4:2	

Note. SD, standard deviation of the mean; cc, cubic centimetres; PPVT, Peabody Picture Vocabulary Test; BNT, Boston Naming Test; SFT, Semantic Fluency Test; COWAT, Controlled Oral Word Association Test; M, Male; F, Female; R, Right; L, Left; A, Ambidextrous.

<sup>1</sup> Thye et al. (2021) did not use any measures derived from diffusion imaging, so 6 participants with only T1 lesion maps and no diffusion data were included in the earlier study but not in this one. One participant with a very large lesion (262.8 cc) was excluded from the earlier study because large lesions are typically uninformative in LSM, but was included in this study because their connectome data may be informative.

**Table 2**  
Diffusion weighted imaging parameters.

Sequence	1	2	3
N	25	18	7
Scanner (3T)	Philips Achieva	Siemens Allegra	Siemens Prisma
Voxel Size (mm)	1.9 × 1.9 × 2.38	2.5×2.5×2.5	2×2×2.6
FOV (mm)	180×161	240×240	244×244
TR (ms)	9403	9400	5500
TE (ms)	69	89	69
Flip angle	90	90	90
Sampling directions	32	32	64
b-Value (s/mm <sup>2</sup> )	800	800	1000

Note. N, number of participants; 3T, 3 T; mm, millimetre; ms, millisecond.

### 2.3. Neuroimaging measures

Automated lesion segmentation using the LINDA package in R (Pustina et al., 2016) was used to distinguish lesioned and spared brain tissue. The lesion masks were binarized and spatially normalized to the same stereotaxic space (Colin27) using symmetric normalization (Avants et al., 2008) as implemented within LINDA. After segmentation, the resulting lesion files were visually examined, and reproducible modifications were made to all lesion masks to account for consistent errors in segmentation (e.g., identifying distal clusters of healthy tissue or portions of the cerebellum as part of the lesion territory). Fig. 1 shows the lesion overlap map.

Whole-brain structural connectomes were generated from each participant's diffusion images using probabilistic tractography. To enable fibre tracking between anatomic atlas regions, participant T1-weighted images were normalized to MNI space using the Clinical Toolbox (Rorden et al., 2012), which accounts for deformations caused by brain lesions, and the inverse transformation matrix was used to transform the AAL3 atlas (Rolls et al., 2020) into native space. The probabilistic grey and white matter tissue maps, lesion mask, and AAL atlas regions were then linearly co-registered to diffusion space using the B0 image as a reference. Where necessary, minor modifications were made to the diffusion pre-processing pipeline (i.e., selection of b0 images) to account for slight differences in scan sequences across participants.

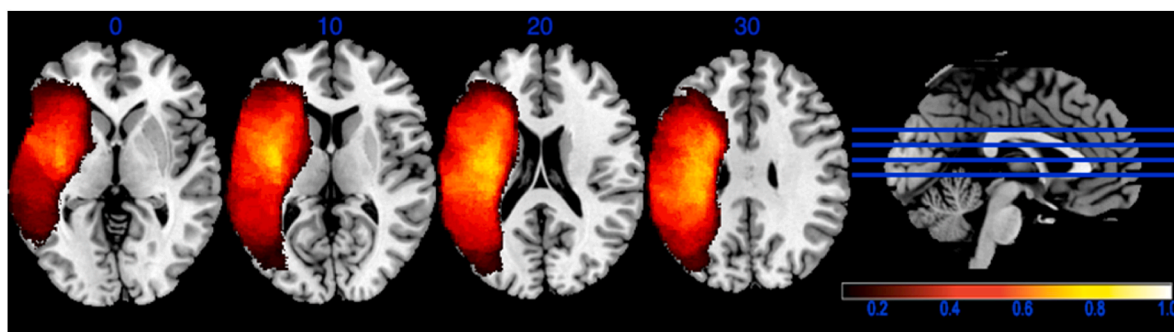
Diffusion data were distortion corrected for motion and eddy currents, and diffusion parameters within each voxel were estimated using FSL FDT's bedpost (Behrens et al., 2007; Jenkinson et al., 2012). Probabilistic fibre tracking was performed in diffusion space using FDT's probtrackX with default tracking parameters (i.e., 5000 individual streamlines drawn in each voxel, 2000 maximum number of steps, 0.5 mm step length, 0.2 curvature threshold, 0.01 fiber volume threshold) and distance correction to quantify the number of streamlines connecting the atlas regions correcting for the distance between the regions. Tractography was restricted to the probabilistic white-matter map excluding the stroke lesion, and streamlines were seeded from each

cortical region. The resulting matrices were corrected for region volume by dividing the number of streamlines between two regions by the combined volume of the connected regions. For each participant, this resulted in a 166×166 symmetric matrix describing the number of pairwise streamlines corrected for the distance between and size of the connected regions.

Connectivity disruption was quantified with graph theoretical measures calculated in MATLAB 2020b (MATLAB, 2020) using the Brain Connectivity Toolbox (Rubinov and Sporns, 2010). For measures of community structure (i.e., transitivity and clustering coefficient), input connectomes were normalized and global network metrics were calculated by taking the mean of the weighted node-level measures. For distance-based measures (i.e., characteristic path length and global efficiency), the normalized connectomes were converted to weighted distance matrices prior to calculation. The measures were generated for a whole-brain network and a left hemisphere language network defined based on prior work using similar methods (Del Gaizo et al., 2017; den Ouden et al., 2019; Fridriksson et al., 2018; Gleichgerrcht et al., 2017; Hula et al., 2020; Marebwa et al., 2017) and consisting of the following regions: middle frontal gyrus, inferior frontal gyrus (pars opercularis, pars triangularis, and pars orbitalis), precentral gyrus, postcentral gyrus, insula, temporal pole (superior and middle portions), inferior temporal gyrus, middle temporal gyrus, superior temporal gyrus, supramarginal gyrus, and angular gyrus.

Four graph theoretical measures were selected for analyses:

- (1) *Average clustering coefficient*: the average of the clustering coefficients of every node within a network. Clustering coefficient is a measure of network segregation, capturing the extent to which functionally related regions are densely connected into specialized clusters. A high clustering coefficient is indicative of a network of densely connected clusters that is robust to damage (i.e., enough connections remain for information to travel to functionally related regions). Damage within the densely connected clusters can impact the efficiency of local processing (Johnson et al., 2020; Kiran et al., 2019).
- (2) *Transitivity*: an alternative to average clustering coefficient for measuring network segregation. Rather than calculating node-level clustering coefficients and averaging them across the network, transitivity is defined at the network level so it is less influenced by individual nodes that have very few connections (Newman, 2003; Rubinov and Sporns, 2010). Transitivity has been found to be associated with better verbal fluency performance in normal aging (Gonzalez-Burgos et al., 2021).
- (3) *Characteristic path length*: the average of the shortest paths between all pairs of nodes within the network. Characteristic path length is a measure of network integration based on how many steps are required to get from one node to another node. Higher values indicate that (on average) information must travel through more steps to get from one node to another, suggesting a less



**Fig. 1.** Lesion overlap map. Only voxels that were lesioned in at least 10% of the participants were included in the analyses. Colours indicate the proportion of participants with lesion in that voxel.

functionally integrated network (Rubinov and Sporns, 2010). It is associated with reduced functional connectivity (Griffis et al., 2020) and the closely related measure of “propagation speed” has been found to be associated with WAB aphasia quotient and fluency (Del Gaizo et al., 2017).

- (4) *Global Efficiency*: the average of each individual node’s global efficiency, which is the average of inverse shortest path lengths between that node and all other nodes. Global efficiency is a measure of network integration that captures how quickly information can travel between brain regions (Latora and Marchiori, 2001). Global efficiency is conceptually closely related to characteristic path length, but the difference in their calculation has substantial implications. If two nodes are not connected, the path length between them is infinite and the communication efficiency is 0. When calculating the average, these values have radically different impact. More generally, characteristic path length is strongly influenced by long paths (infinite path length between disconnected nodes is an extreme case) but efficiency is more sensitive to short paths. This difference may be further exaggerated when these measures are calculated for stroke-damaged brains, which are more likely to have disconnected nodes or to rely on longer paths to communicate around the lesion territory. Johnson et al. (2020) found that higher global efficiency within the semantic network was associated with better response to naming therapy in chronic post-stroke aphasia.

#### 2.4. Statistical analysis

Lesion-symptom mapping analyses were conducted in R (version 3.5.1) (R Development Core Team., 2018) using the LESYMAP package (version 0.0.0.9220) (Pustina, 2019). The R code used for all analyses is provided at <https://osf.io/br3dm>. The analyses were corrected for sufficient lesion involvement by only including voxels that were lesioned in at least 10% of the participants (minimum  $n = 5$ ). All analyses were also corrected for lesion size using direct total lesion volume control (dTLVC) (Mirman et al., 2015). Multivariate LSM was conducted using SCCAN with 4-fold cross-validation to optimise sparseness (Pustina et al., 2018).

Mass-univariate LSM results were corrected for multiple comparisons using continuous permutation-based Family Wise Error Rate (FWER) at  $p < 0.05$ , with  $v = 100$ , meaning that there was a less than 5% chance of observing more than 100 false positive voxels (Mirman et al., 2018). Continuous permutation-based FWER is an extension of traditional permutation-based FWER correction that aligns with the interpretation of VLSM results, which is typically based on groups of voxels (clusters) rather than on single voxels (Mirman et al., 2018). LSM results were labelled using the Automated Anatomical Atlas (AAL) (Tzourio-Mazoyer et al., 2002) and the Johns Hopkins University (JHU) white matter atlas (Hua et al., 2008; Mori et al., 2005).

Connectivity disruption effects were tested using multiple regression to assess different measures of connectivity disruption. Several of the connectivity measures had skewed or bimodal distributions, so the analyses used robust standard error estimation (Savalei, 2014) implemented in the lavaan package version 0.6–8 (Rosseel, 2012). Separate regressions were run for metrics based on the whole brain network and the language sub-network, with overall lesion volume and the five graph theory connectivity metrics entered simultaneously in each regression.

### 3. Results

LSM results are shown in Fig. 2. The mass-univariate VLSM identified 4142 suprathreshold voxels, primarily in the posterior superior temporal lobe (STG, MTG, and Heschl’s gyrus) and a smaller cluster in IFG pars triangularis. Multivariate SCCAN LSM also identified a relatively sparse solution (optimal sparseness = 0.173, CV correlation = 0.55,  $p < 0.0001$ ) with suprathreshold voxels in posterior superior temporal

cortex, IFG pars triangularis, and frontal white matter (superior and anterior corona radiata).

The connectivity disruption regressions revealed only one statistically significant effect: overall lesion volume ( $p < 0.001$  for both analyses). None of the graph theory metrics of connectivity disruption were statistically significant predictors of Token Test performance, regardless of whether they were defined at the whole brain level or for the language sub-network (all  $p > 0.2$ ). See Table 3 for full parameter estimates, standard errors, and confidence intervals.

### 4. Discussion

This study examined the lesion correlates of sentence comprehension deficits in post-stroke aphasia building on prior work on this topic by using a different and clinically-relevant measure of sentence comprehension (the Token Test) and multivariate (SCCAN) and connectome-based lesion-symptom mapping methods. The key findings were that lesions in the posterior superior temporal lobe and IFG pars triangularis were associated with sentence comprehension deficits, which was observed in both mass univariate and multivariate LSM. These regions are often damaged separately (there was only a weak correlation between damage to frontal and temporal clusters:  $r \approx 0.3$ ) and, because stroke lesions tend to be contiguous, damage to both the anterior and posterior clusters would tend to also involve damage to the regions between them. That these intermediate regions were *not* identified by the LSM analysis suggests that the frontal and temporal regions make independent contributions to sentence comprehension deficits. Graph theoretic measures of connectome disruption were not statistically significantly associated with sentence comprehension deficits after accounting for overall lesion size.

These results converge with prior work that identified posterior superior temporal and inferior frontal regions as critical for sentence comprehension (for a review see Wilson, 2017). The Token Test is particularly similar to the sequential commands task and the present LSM results closely replicate the results reported by Lwi et al. (2021), though they did not also find a frontal effect. Lwi et al. also used both mass-univariate and multivariate LSM (finding very good convergence between them), though they used a different multivariate algorithm: SVR-LSM. Unlike SCCAN LSM, SVR-LSM requires correction for multiple comparisons (Sperber et al., 2019), which can produce artifactually focal results (Thye and Mirman, 2018), so that is one possible reason for this difference. It is also possible that the frontal contribution to sentence comprehension is inconsistent across participants. Note that Lwi et al. had a much larger sample size ( $N = 168$ ) than the present study ( $N = 50$ ).

We did not observe involvement of anterior temporal regions, perhaps because the semantic demands were quite limited. Nor did we observe involvement of inferior parietal regions, perhaps because the demands for hierarchical syntactic processing were limited (Matchin and Hickok, 2020). Rather, the critical regions were ones associated with speech perception, verbal WM, and semantic control, suggesting that these are the primary drivers of Token Test performance. The Token Test is also used in clinical contexts, so in addition to providing further insight into the neural correlates of sentence comprehension deficits, these results are relevant for interpreting Token Test results in clinical settings.

In sum, the present study examined auditory sentence comprehension deficits in post-stroke aphasia using new methods: a clinically-relevant behavioural measure, multivariate LSM, and connectome disruption measures. The results largely converge with prior work, highlighting the critical role of posterior superior temporal and inferior frontal regions in sentence comprehension, and providing a valuable (and rare) replication of those effects. The results also suggest more specialized roles for anterior temporal and inferior parietal regions as well as caution regarding use of connectome disruption measures.

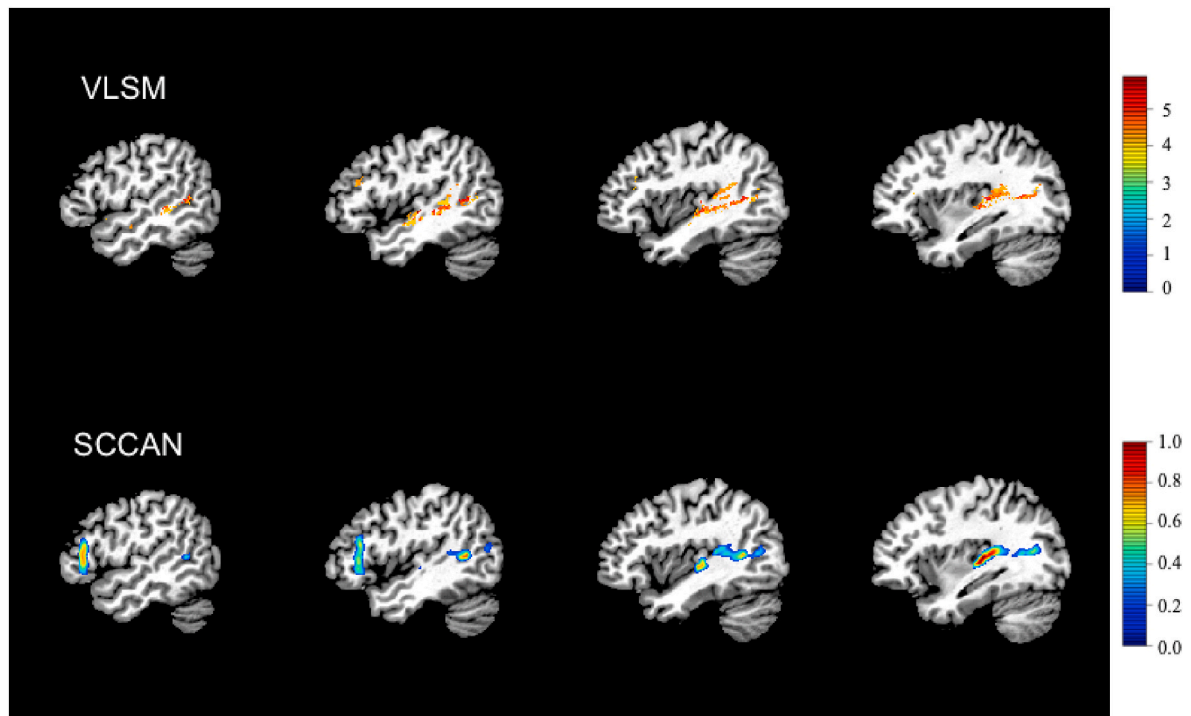


Fig. 2. Lesion-symptom mapping (LSM) results. Top row shows suprathreshold voxels from mass-univariate VLSM results. Bottom row shows SCCAN LSM solution.

**Table 3**

Results of connectivity disruption regression analyses. Values are the coefficient estimates with standard errors in parentheses and 95% confidence intervals in brackets. Only lesion volume was statistically significantly associated with Token Test performance.

Metric	Whole Brain Network	Language Network
Lesion volume (cc)	-0.111 (0.024), [-0.158, -0.064]	-0.116 (0.03), [-0.174, -0.057]
Average Clustering Coefficient	6811 (6773), [-6463, 20085]	-20.7 (156), [-327, 286]
Transitivity	-4616 (6433), [-17224, 7992]	133 (136), [-134, 400]
Characteristic path length	0.011 (0.021), [-0.031, 0.053]	0.001 (0.003), [-0.004, 0.007]
Global efficiency	-202.4 (187.2), [-569.3, 164.5]	-14.9 (34.8), [-83.1, 53.2]

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgements

The collection of data analysed here was supported by grants NIH R01 HD068488 and R01 NS048281 (PI: Szaflarski); data analyses were supported by NIH R01DC017137 (MPI: Mirman/Szaflarski). This research was also funded, in part, by the Wellcome Trust [108890/Z/15/Z]. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

### References

Allendorfer, J.B., Kissela, B.M., Holland, S.K., Szaflarski, J.P., 2012. Different patterns of language activation in post-stroke aphasia are detected by overt and covert versions

of the verb generation fMRI task. *Medical Science Monitor* 18 (3), CR135–CR147. <https://doi.org/10.12659/MSM.882518>.

- Allendorfer, J.B., Nenert, R., Nair, S., Vannest, J., Szaflarski, J.P., 2021. Functional resonance imaging of language following constraint-induced Aphasia therapy primed with intermittent theta burst stimulation in 13 patients with post-stroke Aphasia. *Med. Sci. Mon. Int. Med. J. Exp. Clin. Res.* 27 <https://doi.org/10.12659/MSM.930100>.
- Avants, B.B., Epstein, C.L., Grossman, M., Gee, J.C., 2008. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* 12 (1), 26–41. <https://doi.org/10.1016/j.media.2007.06.004>.
- Barbey, A.K., Colom, R., Grafman, J., 2014. Neural mechanisms of discourse comprehension: a human lesion study. *Brain* 137 (1), 277–287. <https://doi.org/10.1093/brain/awt312>.
- 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 (5), 448–450. <https://doi.org/10.1038/nn1050>.
- Behrens, T.E.J., Berg, H.J., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W., 2007. Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *Neuroimage* 34 (1), 144–155. <https://doi.org/10.1016/j.neuroimage.2006.09.018>.
- Boebel, W., Wagenmakers, E.-J., Belay, L., Verhagen, J., Brown, S., Forstmann, B.U., 2015. A purely confirmatory replication study of structural brain-behavior correlations. *Cortex* 66, 115–133. <https://doi.org/10.1016/j.cortex.2014.11.019>.
- De Renzi, E., Vignolo, L.A., 1962. The token test: a sensitive test to detect receptive disturbances in aphasics. *Brain* 85 (4), 665–678. <https://doi.org/10.1093/brain/85.4.665>.
- Del Gaizo, J., Fridriksson, J., Yourganov, G., Hillis, A.E., Hickok, G., Misis, B., Rorden, C., Bonilha, L., 2017. Mapping language networks using the structural and dynamic brain connectomes. *Eneuro* 4 (5), 204–217. <https://doi.org/10.1523/ENEURO.0204-17.2017>. ENEURO, 2017.
- den Ouden, D.-B., Malyutina, S., Basilakos, A., Bonilha, L., Gleichgerrcht, E., Yourganov, G., Hillis, A.E., Hickok, G., Rorden, C., Fridriksson, J., 2019. Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. *Hum. Brain Mapp.* 40 (7), 2153–2173. <https://doi.org/10.1002/hbm.24514>.
- 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 (1–2), 145–177. <https://doi.org/10.1016/j.cognition.2003.11.002>.
- Fridriksson, J., den Ouden, D.-B., Hillis, A.E., Hickok, G., Rorden, C., Basilakos, A., Yourganov, G., Bonilha, L., 2018. Anatomy of aphasia revisited. *Brain* 141 (3), 848–862. <https://doi.org/10.1093/brain/awx363>.
- Geller, J., Thyne, M., Mirman, D., 2019. Estimating effects of graded white matter damage and binary tract disconnection on post-stroke language impairment. *Neuroimage* 189, 248–257. <https://doi.org/10.1016/j.neuroimage.2019.01.020>.
- Gleichgerrcht, E., Fridriksson, J., Rorden, C., Bonilha, L., 2017. Connectome-based lesion-symptom mapping (CLSM): a novel approach to map neurological function.

- Neuroimage: Clinical 16 (July), 461–467. <https://doi.org/10.1016/j.nicl.2017.08.018>.
- Gonzalez-Burgos, L., Pereira, J.B., Mohanty, R., Barroso, J., Westman, E., Ferreira, D., 2021. Cortical networks underpinning compensation of verbal fluency in normal aging. *Cerebr. Cortex* 31 (8), 3832–3845. <https://doi.org/10.1093/cercor/bhab052>.
- Griffis, J.C., Metcalfe, N.V., Corbetta, M., Shulman, G.L., 2020. Damage to the shortest structural paths between brain regions is associated with disruptions of resting-state functional connectivity after stroke. *Neuroimage* 210 (116589). <https://doi.org/10.1016/j.neuroimage.2020.116589>.
- Griffis, J.C., Nenert, R., Allendorfer, J.B., Vannest, J., Holland, S., Dietz, A., Szaflarski, J.P., 2017. The canonical semantic network supports residual language function in chronic post-stroke aphasia. *Hum. Brain Mapp.* 38 (3), 1636–1658. <https://doi.org/10.1002/hbm.23476>.
- Grodzinsky, Y., Santi, A., 2008. The battle for Broca's region. *Trends Cognit. Sci.* 12 (12), 474–480. <https://doi.org/10.1016/j.tics.2008.09.001>.
- Hickok, G.S., Poeppel, D., 2007. The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8 (May), 393–402.
- 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 (1), 336–347. <https://doi.org/10.1016/j.neuroimage.2007.07.053>.
- Hula, W.D., Panesar, S., Gravier, M.L., Yeh, F.-C., Dresang, H.C., Dickey, M.W., Fernandez-Miranda, J.C., 2020. Structural white matter connectometry of word production in aphasia: an observational study. *Brain* 143 (8), 2532–2544. <https://doi.org/10.1093/brain/awaa193>.
- Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. *Neuroimage* 62 (2), 782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>.
- Johnson, J.P., Meier, E.L., Pan, Y., Kiran, S., 2020. Pre-treatment graph measures of a functional semantic network are associated with naming therapy outcomes in chronic aphasia. *Brain Lang.* 207, 104809. <https://doi.org/10.1016/j.bandl.2020.104809>.
- Kiran, S., Meier, E.L., Johnson, J.P., 2019. Neuroplasticity in Aphasia: a proposed framework of language recovery. *J. Speech Lang. Hear. Res.* 62 (11), 3973–3985. [https://doi.org/10.1044/2019\\_JSLHR-L-RSNP-19-0054](https://doi.org/10.1044/2019_JSLHR-L-RSNP-19-0054).
- Kristinsson, S., Thors, H., Yourganov, G., Magnusdottir, S., Hjaltason, H., Stark, B.C., Basilakos, A., den Ouden, D.-B., Bonilha, L., Rorden, C., Hickok, G., Hillis, A., Fridriksson, J., 2020. Brain damage associated with impaired sentence processing in acute Aphasia. *J. Cognit. Neurosci.* 32 (2), 256–271. [https://doi.org/10.1162/jocn\\_a.01478](https://doi.org/10.1162/jocn_a.01478).
- Latora, V., Marchiori, M., 2001. Efficient behavior of small-World networks. *Phys. Rev. Lett.* 87 (19), 198701. <https://doi.org/10.1103/PhysRevLett.87.198701>.
- Leff, A.P., Schofield, T.M., Crinion, J.T., Seghier, M.L., Grogan, A., Green, D.W., Price, C.J., 2009. The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: evidence from 210 patients with stroke. *Brain* 132 (12), 3401–3410. <https://doi.org/10.1093/brain/awp273>.
- Lwi, S.J., Herron, T.J., Curran, B.C., Ivanova, M.V., Schendel, K., Dronkers, N.F., Baldo, J.V., 2021. Auditory comprehension deficits in post-stroke Aphasia: Neurologic and demographic correlates of outcome and recovery. *Front. Neurol.* 12 (1323) <https://doi.org/10.3389/fneur.2021.680248>.
- Marebwa, B.K., Fridriksson, J., Yourganov, G., Feenaughty, L., Rorden, C., Bonilha, L., 2017. Chronic post-stroke aphasia severity is determined by fragmentation of residual white matter networks. *Scientific Reports* 7 (1), 8188. <https://doi.org/10.1038/s41598-017-07607-9>.
- Matchin, W., Hickok, G., 2020. The cortical organization of syntax. *Cerebr. Cortex* 30 (3), 1481–1498. <https://doi.org/10.1093/cercor/bhz180>.
- MATLAB (9.9.0 (R2020b)), 2020. [Computer software]. The Mathworks Inc.
- Mirman, D., Britt, A.E., 2014. What we talk about when we talk about access deficits. *Phil. Trans. Biol. Sci.* 369 (1634), 1–14. <https://doi.org/10.1098/rstb.2012.0388>.
- Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O.K., Coslett, H.B., Schwartz, M.F., 2015. Neural organization of spoken language revealed by lesion-symptom mapping. *Nat. Commun.* 6 (6762), 1–9. <https://doi.org/10.1038/ncomms7762>.
- Mirman, D., Landrigan, J.-F., Kokolis, S., Verillo, S., Ferrara, C., Pustina, D., 2018. Corrections for multiple comparisons in voxel-based lesion-symptom mapping. *Neuropsychologia* 115, 112–123. <https://doi.org/10.1016/j.neuropsychologia.2017.08.025>.
- Mori, S., Wakana, S., van Zijl, P.C.M., Nagae-Poetscher, L.M., 2005. *MRI Atlas of Human White Matter*. Elsevier.
- Nenert, R., Allendorfer, J.B., Martin, A.M., Banks, C., Ball, A., Vannest, J., Dietz, A.R., Szaflarski, J.P., 2017. Neuroimaging correlates of post-stroke Aphasia rehabilitation in a pilot randomized trial of constraint-induced Aphasia therapy. *Med. Sci. Monit.* 23, 3489–3507. <https://doi.org/10.12659/MSM.902301>.
- Newhart, M., Trupe, L.A., Gomez, Y., Cloutman, L., Molitoris, J.J., Davis, C., Leigh, R., Gottesman, R.F., Race, D., Hillis, A.E., 2012. Asyntactic comprehension, working memory, and acute ischemia in Broca's area versus angular gyrus. *Cortex* 48 (10), 1288–1297. <https://doi.org/10.1016/j.cortex.2011.09.009>.
- Newman, M.E.J., 2003. The structure and function of complex networks. *SIAM Rev.* 45 (2), 167–256. <https://doi.org/10.1137/S003614450342480>.
- Pillay, S.B., Binder, J.R., Humphries, C., Gross, W.L., Book, D.S., 2017. Lesion localization of speech comprehension deficits in chronic aphasia. *Neurology* 88 (10), 970–975. <https://doi.org/10.1212/WNL.0000000000003683>.
- Pustina, D., 2019. *LESYMAP: Lesion to Symptom Mapping in R* (0.0.0.9220) [Computer software].
- Pustina, D., Avants, B., Faseyitan, O.K., Medaglia, J.D., Coslett, H.B., 2018. Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. *Neuropsychologia* 115, 154–166. <https://doi.org/10.1016/j.neuropsychologia.2017.08.027>.
- Pustina, D., Coslett, H.B., Turkeltaub, P.E., Tustison, N., Schwartz, M.F., Avants, B.B., 2016. Automated segmentation of chronic stroke lesions using LINDA: lesion identification with neighborhood data analysis. *Hum. Brain Mapp.* 37 (4), 1405–1421. <https://doi.org/10.1002/hbm.23110>.
- R Development Core Team, 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.r-project.org/>.
- Rogalsky, C., LaCroix, A.N., Chen, K.-H., Anderson, S.W., Damasio, H., Love, T., Hickok, G., 2018. The Neurobiology of agrammatic sentence comprehension: a lesion study. *J. Cognit. Neurosci.* 30 (2), 234–255. [https://doi.org/10.1162/jocn\\_a.01200](https://doi.org/10.1162/jocn_a.01200).
- Rolls, E.T., Huang, C.-C., Lin, C.-P., Feng, J., Joliot, M., 2020. Automated anatomical labelling atlas 3. *Neuroimage* 206 (116189). <https://doi.org/10.1016/j.neuroimage.2019.116189>.
- Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., Karnath, H.-O., 2012. Age-specific CT and MRI templates for spatial normalization. *Neuroimage* 61 (4), 957–965. <https://doi.org/10.1016/j.neuroimage.2012.03.020>.
- Rossee, Y., 2012. lavaan: an R package for structural equation modeling. *J. Stat. Software* 48 (1), 1–36. <https://doi.org/10.18637/jss.v048.i02>.
- Rubinow, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52 (3), 1059–1069. <https://doi.org/10.1016/j.neuroimage.2009.10.003>.
- Savalei, V., 2014. Understanding robust corrections in structural equation modeling. *Struct. Equ. Model.* 21 (1), 149–160. <https://doi.org/10.1080/10705511.2013.824793>.
- Sperber, C., Wiesen, D., Karnath, H.-O., 2019. An empirical evaluation of multivariate lesion behaviour mapping using support vector regression. *Hum. Brain Mapp.* 40, 1381–1390. <https://doi.org/10.1002/hbm.24476>.
- Szaflarski, J.P., Ball, A.L., Vannest, J., Dietz, A.R., Allendorfer, J.B., Martin, A.N., Hart, K., Lindsell, C.J., 2015. Constraint-induced Aphasia therapy for treatment of chronic post-stroke Aphasia: a randomized, blinded, controlled pilot trial. *Med. Sci. Monit.* 21, 2861–2869. <https://doi.org/10.12659/MSM.894291>.
- Szaflarski, J.P., Nenert, R., Allendorfer, J.B., Martin, A.N., Amara, A.W., Griffis, J.C., Dietz, A., Mark, V.W., Sung, V.W., Walker, H.C., Zhou, X., Lindsell, C.J., 2021. Intermittent theta burst stimulation (iTBS) for treatment of chronic post-stroke Aphasia: results of a pilot randomized, double-blind, sham-controlled trial. *Med. Sci. Monit.* 27 <https://doi.org/10.12659/MSM.931468>.
- Thothathiri, M., Kimberg, D.Y., Schwartz, M.F., 2012. The neural basis of reversible sentence comprehension: evidence from voxel-based lesion symptom mapping in Aphasia. *J. Cognit. Neurosci.* 24 (1), 212–222. [https://doi.org/10.1162/jocn\\_a.00118](https://doi.org/10.1162/jocn_a.00118).
- Thye, M., Mirman, D., 2018. Relative contributions of lesion location and lesion size to predictions of varied language deficits in post-stroke aphasia. *Neuroimage: Clinical* 20 (October), 1129–1138. <https://doi.org/10.1016/j.nicl.2018.10.017>.
- Thye, M., Szaflarski, J.P., Mirman, D., 2021. Shared lesion correlates of semantic and letter fluency in post-stroke aphasia. *J. Neuropsychol.* 15 (1), 143–150. <https://doi.org/10.1111/jnp.12211>.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15 (1), 273–289. <https://doi.org/10.1006/nimg.2001.0978>.
- Wilson, S.M., 2017. Lesion-symptom mapping in the study of spoken language understanding. *Language, Cognition and Neuroscience* 32 (7), 891–899. <https://doi.org/10.1080/23273798.2016.1248984>.