

White matter microstructural integrity pre- and post-treatment in individuals with chronic post-stroke aphasia

Emily J. Braun^{a,*}, Anne Billot^{a,b}, Erin L. Meier^{a,1,3}, Yue Pan^{a,2,3}, Todd B. Parrish^c, Ajay S. Kurani^d, Swathi Kiran^a

^a Aphasia Research Laboratory, Department of Speech, Language & Hearing Sciences, College of Health & Rehabilitation Sciences, Sargent College, Boston University, 635 Commonwealth Avenue, Room 326, Boston, MA 02115, USA

^b School of Medicine, Boston University, Boston, MA, USA

^c Department of Radiology, Feinberg School of Medicine, Northwestern University, 737 N. Michigan Avenue, Suite 1600, Chicago, IL 60611, USA

^d Department of Neurology, Feinberg School of Medicine, Northwestern University, 625 N. Michigan Avenue, Suite 1150, Chicago, IL 60611, USA

ARTICLE INFO

Keywords:

Aphasia
Stroke
White matter
Diffusion tensor imaging
Rehabilitation
Automated fiber quantification
Tractography
Language

ABSTRACT

While previous studies have found that white matter damage relates to impairment severity in individuals with aphasia, further study is required to understand the relationship between white matter integrity and treatment response. In this study, 34 individuals with chronic post-stroke aphasia underwent behavioral testing and structural magnetic resonance imaging at two timepoints. Thirty participants within this sample completed typicality-based semantic feature treatment for anomia. Tractography of bi-hemispheric white matter tracts was completed via Automated Fiber Quantification. Associations between microstructural integrity metrics and behavioral measures were evaluated at the tract level and in nodes along the tract. Diffusion measures of the left inferior longitudinal, superior longitudinal, and arcuate fasciculi were related to aphasia severity and diffusion measures of the left inferior longitudinal fasciculus were related to naming and treatment response. This study also found preliminary evidence of left inferior longitudinal fasciculus microstructural changes following treatment.

1. Introduction

Stroke is a leading cause of disability. In the United States alone, more than 7.6 million adults have had a stroke (Virani et al., 2021) and approximately 30 % of ischemic strokes or mixed ischemic-hemorrhagic strokes result in aphasia (Flowers et al., 2016). Aphasia affects social relationships and life participation. Understanding the neural underpinnings of aphasia provides insight into the underlying nature of impairment and ultimately how best to achieve optimal rehabilitation outcomes. In particular, individuals with aphasia often present with gray and white matter damage which both contribute to impairment.

Identifying white matter tracts (hereafter also referred to as “tracts”) implicated in language processing in neurotypical individuals is key to

understanding how structural damage relates to language processing in individuals with aphasia. Language processing in healthy individuals is structurally and functionally subserved by a left-lateralized network of frontal, temporal, and parietal gray matter cortical regions connected by white matter tracts (Binder et al., 2009; Friederici, 2002, 2011, 2012; Price, 2010, 2012; Turken & Dronkers, 2011). The dual-stream model of speech processing provides a framework for understanding this network (Hickok & Poeppel, 2004, 2007). In the dual-stream model, through left-lateralized dorsal stream processes, phonological representations are mapped to articulatory motor representations (Hickok & Poeppel, 2004, 2007). Along with gray matter cortical regions, white matter fiber bundles of the superior longitudinal fasciculus (SLF) and arcuate fasciculus (AF) support these processes (Saur et al., 2008). In

Abbreviations: CC major, corpus callosum forceps major; CC minor, corpus callosum forceps minor; AF, arcuate fasciculus; SLF, superior longitudinal fasciculus; ILF, inferior longitudinal fasciculus; IFOF, inferior fronto-occipital fasciculus; UF, uncinate fasciculus; FA, fractional anisotropy; MD, mean diffusivity.

* Corresponding author at: 635 Commonwealth Avenue, Room 326, Boston, MA 02215, USA.

E-mail address: ejbraun@bu.edu (E.J. Braun).

¹ Present affiliation: Department of Communication Sciences and Disorders, Northeastern University, Boston, MA, USA.

² Present affiliation: Dorothy M. Davis Heart and Lung Research Institute, The Ohio State University, Columbus, OH, USA.

³ Note: Erin L. Meier and Yue Pan were at Boston University at the time this work was initiated.

<https://doi.org/10.1016/j.bandl.2022.105163>

Received 27 July 2021; Received in revised form 21 July 2022; Accepted 23 July 2022

Available online 31 July 2022

0093-934X/© 2022 Elsevier Inc. All rights reserved.

complement, through more bilateral ventral stream processes, phonological representations are mapped to lexical-semantic representations (Hickok & Poeppel, 2004, 2007). White matter fiber bundles of the inferior longitudinal fasciculus (ILF), inferior fronto-occipital fasciculus (IFOF), and potentially the uncinate fasciculus (UF) support these processes (Saur et al., 2008; Turken & Dronkers, 2011). The corpus callosum (CC) connects the two hemispheres and contributes to a number of language functions (Friederici et al., 2007). This knowledge of neurotypical structural connectivity provides a guide for white matter regions to examine in individuals with aphasia.

In individuals with post-stroke aphasia, converging evidence has shown reduced macrostructural (i.e., related to macroscopic anatomy) and microstructural (i.e., related to microscopic anatomy) integrity of white matter tracts important for language (Mori, 2009). Importantly, in terms of macrostructural integrity, individuals with aphasia exhibit reduced number of fibers (van Hees et al., 2014) and in some cases an inability to delineate left-hemisphere tracts as a result of the stroke lesion (Kim & Jang, 2013; Tak & Jang, 2014). This inability to delineate tracts is thought to be due to more severe damage or degeneration and has been associated with more severe language impairment (Tak & Jang, 2014).

Additionally, multiple studies have shown reduced microstructural integrity of left-hemisphere tracts compared to right-hemisphere homologues and left-hemisphere tracts in neurotypical controls (Ivanova et al., 2016; McKinnon et al., 2017; van Hees et al., 2014). These studies have examined several microstructural integrity metrics given different underlying mechanisms represented by each (Alexander et al., 2008; Tae et al., 2018). Two commonly used measures, fractional anisotropy (FA) and mean diffusivity (MD), index changes thought to occur post-stroke. FA measures restriction of water molecule diffusion (Pierpaoli et al., 1996) and is considered to be a marker of axonal integrity (Geva et al., 2011; Tae et al., 2018). In general, higher FA values are indicative of greater microstructural integrity. MD gives information related to average magnitude of water displacement and is thought to be particularly sensitive to edema and necrosis (Campbell & Pike, 2014; Soares et al., 2013; Tae et al., 2018). In general, lower MD values are indicative of greater microstructural integrity.

This reduced macrostructural (e.g., lesion load, number of fibers, presence or absence of delineated tracts) and microstructural (e.g., DTI scalars such as FA and MD values) integrity relates to behavioral performance. Table 1 shows that white matter integrity has been associated with different aspects of language performance in dorsal and ventral left-hemisphere structures, including the AF (Hula et al., 2020; Ivanova et al., 2016; Keser, Meier, Stockbridge, & Hillis, 2020a; Kim & Jang, 2013; J.-K. Lee et al., 2021; S. Lee et al., 2020; Tak & Jang, 2014; Zhang et al., 2021), SLF (Han et al., 2016; J.-K. Lee et al., 2021; McKinnon et al., 2018; Ramsey et al., 2017; Yang et al., 2017), ILF (Harvey & Schnur, 2015; Ivanova et al., 2016; Keser, Meier, Stockbridge, & Hillis, 2020a; McKinnon et al., 2018; Meier, Johnson, Pan, & Kiran, 2019b; Xing, Lacey, Skipper-Kallal, Zeng, & Turkeltaub, 2017; Zhang et al., 2018, 2021), IFOF (Harvey & Schnur, 2015; Hula et al., 2020; Ivanova et al., 2016; J.-K. Lee et al., 2021; Meier et al., 2019b; Xing et al., 2017; Yang et al., 2017; Zhang et al., 2018, 2021) and UF (Hula et al., 2020; Xing et al., 2017; Zhang et al., 2018, 2021). Limited studies have shown a correlation between diffusion metrics in the right hemisphere and language in individuals with aphasia (Forkel et al., 2014; Hartwigsen et al., 2020; Osa García et al., 2020; Pani et al., 2016).

In addition to the relationship with single-timepoint language skills, white matter structural integrity metrics have also been correlated to aphasia treatment gains. Overall, results from studies evaluating behavioral recovery over time suggest that pre-treatment or initial timepoint left-hemisphere white matter integrity relates positively to treatment response or language outcomes (Bonilha, Gleichgerrcht, Nesland, Rorden, & Fridriksson, 2016; Keser, Meier, Stockbridge, & Hillis, 2020a; Meier, Johnson, Pan, & Kiran, 2019b; van Hees et al., 2014). In addition, initial reports evaluating white matter plasticity

following behavioral treatment in individuals with chronic post-stroke aphasia suggest that bilateral tracts undergo changes during recovery and that changes in white matter correlate with behavioral changes (Blom-Smink et al., 2020; Jang et al., 2017; Keser, Sebastian, et al., 2020; McKinnon et al., 2017; Schlaug et al., 2009; van Hees et al., 2014; Wan et al., 2014). For example, McKinnon and colleagues found a relationship between change in mean kurtosis (MK⁴) of the left ILF and reduction in semantic paraphasias in individuals who completed intensive language action therapy (ILAT) when evaluating subsections of the ILF (McKinnon et al., 2017). Because white matter tracts traverse larger areas of the brain, different areas of the tract may show different levels of microstructural abnormality, as shown by McKinnon and colleagues. Potential reasons for along-tract differences in microstructural abnormality include varying susceptibility to damage and varying distance from the lesion. Overall, further investigation is required given the limited number of studies with behavioral and neural data at two timepoints. Detailed examination of macrostructural and microstructural tract properties, including subsections of tracts, will better characterize fine-grained post-stroke changes in microstructural integrity and provide clarity to variable results in the literature.

As Table 1 shows, there is notable variability in methodology and results across studies examining relationships between white matter metrics and language in aphasia. Options such as choice of method (e.g., tractography vs atlas-based methods); choice of approach within the method (e.g., deterministic vs probabilistic tractography); parameters; which portions of tracts to evaluate; and microstructural integrity metric impact results and should be considered as a source of variability when interpreting inter-study findings.

The current study examined the relationship between microstructural integrity of specific white matter tracts (as indexed by FA and MD values in whole tracts and tract segments) and naming skills, aphasia severity, and treatment response. Particular attention was given to documenting the number of tracts delineated at two selected thresholds. A crucial assumption of this paper was that careful tract delineation would reveal nuanced relationships between integrity of specific tracts and language performance. This study evaluated the same group of participants analyzed by Meier and colleagues (Meier et al., 2019b), with the following analytical choices: (a) use of pre- and post-treatment longitudinal neuroimaging data, (b) use of tractography, (c) evaluation of white matter tract subsegments, and (d) inclusion of multiple DTI scalars. Our research questions were:

1. Question 1: In individuals with chronic post-stroke aphasia, what is the relationship between white matter tract integrity and language skills? Specifically:
 - a. What is the relationship between macrostructural integrity (i.e., white matter tract delineation) and aphasia severity? We expected to find that presence of tract delineation in the left hemisphere would be related to better language performance.
 - b. Does white matter microstructural integrity (i.e., FA and MD values) relate to aphasia severity? We hypothesized that greater white matter tract microstructural integrity of left-hemisphere dorsal and ventral delineated tracts would be related to milder aphasia at Timepoint 1.
 - c. Does white matter microstructural integrity (i.e., FA and MD values) relate to naming ability? We hypothesized that greater white matter tract microstructural integrity of left-hemisphere dorsal and ventral delineated tracts would be related to greater naming ability at Timepoint 1.
2. Question 2: Does white matter microstructural integrity (i.e., FA and MD values) relate to naming treatment response? We hypothesized that greater white matter tract microstructural integrity of left-

⁴ Mean kurtosis is thought to index microstructural complexity (Lätt et al., 2013).

Table 1
Studies correlating white matter metrics with language in individuals with aphasia.

1 st Author	Year	Sample Size	Method	Metric	other	Left				Right	
						AF/SLF	ILF	IFOF	UF	AF/SLF	ILF
Dresang	2021	14	connectometry	connectome matrix	✓ ²	✓	✓	✓			
Hula	2020	42	connectometry	connectome matrix	✓ ³	✓		✓	✓		
Lee, J.-K.	2021	64	tractography	FA		✓		✓			
Blom-Smink	2020	10	tractography	FA							*
Lee, S.	2020	68	tractography	LI-AD, LI-FD, delineation		✓					
Osa García	2020	20	tractography	FA						✓	
Yu ¹	2019	20	tractography	delineation							
McKinnon	2018	32	tractography	AWF		✓	✓				
Zhang	2018	14	tractography; TBSS	FA			✓	✓	✓		
McKinnon	2017	8	tractography	MK			*				
Jang	2017	16	tractography	volume		*					
Harvey & Schnur	2015	15	tractography	FA			✓	✓			
Bonilha	2014	8	tractography; connectome	# of lesioned fibers	✓ ⁴						
van Hees	2014	8	tractography	GFA		*					
Forkel	2014	16	tractography	volume						✓	
Tak & Jang	2014	25	tractography	volume, delineation		✓					
Kim & Jang	2013	25	tractography	delineation		✓					
Harvey	2013	10	tractography	FA					✓		
Hosomi	2009	13	tractography	# of fibers		✓ ⁵					
Schlaug	2009	6	tractography	# of fibers						* ⁶	
Hartwigsen	2020	12	TBSS	FA						✓	
Yang	2017	18	TBSS	FA		✓		✓			
Xing	2017	40	TBSS	FA, MD, AD, RD			✓	✓	✓		
Geva	2015	15	tractography; atlas-based; TBSS	delineation		✓					
Keser	2020b	24	atlas-based	FA						* ⁷	
Keser	2020a	28	atlas-based	FA, RD		✓	✓	✓			
Meier	2019b	34	atlas-based	FA			✓*	✓*			
Han	2016	69	atlas-based	FA, lesion load		✓					
Ivanova	2016	37	atlas-based	FA		✓	✓	✓			
Pani	2016	33	atlas-based	FA, lesion load	✓ ⁸	✓					
Rosso	2015	23	atlas-based	FA		✓		✓			
Wan	2014	11	atlas-based	FA	* ⁹						
Han	2013	83	atlas-based	FA	✓ ¹⁰			✓	✓		
Papoutsis	2011	14	atlas-based	FA, MD	✓ ¹¹	✓					
Rolheiser	2011	24	atlas-based	FA	✓ ¹¹	✓					
Hillis	2018	19 + 159	atlas-based	lesion load		✓*					
Ramsey	2017	132 ¹²	atlas-based	lesion load		✓					
Kümmeler	2013	100	atlas-based	lesion load	✓ ¹³	✓					
Wang	2013	50	atlas-based	lesion load		✓					
Marchina	2011	30	atlas-based	lesion load		✓					
Meinzer	2010	10	individual ROIs	FA	* ¹⁴						

✓white matter (WM) measures correlated with language task performance; *WM measures or change in WM measures correlated with change in behavioral measures; **Abbreviations:** axial diffusivity (AD), radial diffusivity (RD), fiber density (FD), Laterality index (LI), mean kurtosis (MK), generalized FA (GFA), tract-based spatial statistics (TBSS), left hemisphere (LH); ¹no statistically significant results; ²cortico-subcortical projection pathways; ³limbic pathways, middle longitudinal fasciculi; ⁴left cingulum (hippocampus), exploratory analysis; ⁵loss of leftward asymmetry; ⁶trend toward correlation between change in AF number of fibers and change in CIUs/minute; ⁷negative correlation between changes in right arcuate fasciculus and recovery rate in naming; ⁸white matter underlying right MTG, right precentral gyrus, and right IFG pars opercularis; corpus callosum connecting left and right supplementary motor areas; ⁹improvements in CIUs/minute correlated with reductions in FA in WM underlying right IFG; ¹⁰anterior thalamic radiation; ¹¹extreme capsule; ¹²40 participants excluded for various reasons; ¹³composite left-hemisphere dorsal and ventral streams; ¹⁴hippocampus adjacent white matter.

atlas-based: either (a) information from structural scans (typically diffusion-weighted) overlaid onto atlases of canonical white matter tracts or (b) lesion overlay approach (lesion maps intersected with canonical tracts to determine how percentage of tract damage (i.e., lesion load) relates to language function).

hemisphere dorsal and ventral delineated tracts at pre-treatment would be related to more favorable behavioral response to anomia treatment.

- Question 3: Does white matter microstructural integrity (i.e., FA and MD values) change following naming treatment? We hypothesized that delineated left-hemisphere dorsal and ventral tracts would show increased microstructural integrity from pre- to post-treatment in individuals who completed anomia treatment. Successful treatment completion engages semantic and phonological processing necessary for retrieval and production and, as such, we expected both dorsal and ventral tracts to show changes.

We did not expect right-hemisphere tract microstructural integrity to be associated with language performance or treatment gains.

2. Methods

2.1. Participants

Participants were 34 individuals (24 males) with chronic post-stroke aphasia with mean age 62 years (SD = 10.8) and mean time post-stroke onset 62 months (SD = 86) (see Supplementary Table 1 for individual demographic data). Exclusion criteria were premonitory neurological disease, history of multiple left-hemisphere strokes, and contraindications for MRI. Written informed consent was obtained from all participants prior to participation. The study was approved by the institutional review boards of Boston University, Massachusetts General Hospital, and Northwestern University. Thirty-two participants completed the study through the Aphasia Research Laboratory at Boston University and two participants completed the study through the Aphasia and Neurolinguistics Research Laboratory at Northwestern University. The

study included 34 unique participants, of which 30 participants received treatment and 11 served as no-treatment natural history control participants.⁵ Previous publications reported behavioral (Gilmore et al., 2020) and neuroimaging (Johnson et al., 2019, 2020; Meier et al., 2018, 2019a, 2019b) findings for this group of participants.

2.2. Behavioral assessment and treatment

Participants completed standardized assessment of aphasia severity via the Western Aphasia Battery – Revised Aphasia Quotient (WAB-R AQ) (Kertesz, 2007) and naming via the Boston Naming Test – Second Edition (BNT-2) (Goodglass et al., 2001). This testing was completed at two timepoints approximately 12 weeks apart, corresponding to pre- and post-treatment for participants who underwent treatment. Treated participants also completed study-specific confrontation naming probes for 36 trained items (18 items each from two semantic categories out of *birds*, *vegetables*, *clothing*, and *furniture*) at pre- and post-treatment timepoints. Treated participants ($N = 30$) completed up to 24 two-hour sessions of typicality-based semantic feature treatment with the goal of improved naming through targeting semantic processing. In this treatment, intervention included steps aimed to strengthen lexical-semantic representations (i.e., sorting items by category, verifying whether particular semantic features apply to a noun, and confrontation naming) (Gilmore et al., 2020). Some participants were trained with atypical category exemplars and generalization to more typical category exemplars was assessed. This approach is based on the Complexity Account of Treatment Efficacy, which asserts that training of more complex items will generalize to related less complex items (Kiran & Thompson, 2003). Treatment response was measured by proportion of potential maximal gain (PMG) for trained items (post-treatment naming score – baseline naming score)/(number of trained items – baseline naming score) (Gilmore et al., 2019; Lambon Ralph et al., 2010). This approach allows for calculating treatment gains while taking into account pre-treatment performance. When possible, each treatment group participant also served as a natural history control (i.e., two baselines taken over a period of approximately 12 weeks were collected prior to treatment onset).

2.3. Image acquisition

Participants underwent MR imaging on either a Siemens 3 T Skyra at the Athinoula A. Martinos Center in Charlestown, MA ($n = 32$) or on a Siemens 3 T Prisma Fit at the Center for Translational Imaging in Chicago, IL ($n = 2$).⁶ T1-weighted sagittal imaging was collected (TR/TE = 2300/2.98 ms, TI = 900 ms, flip angle = 9° , FOV = 256×256 mm², voxel size = $1 \times 1 \times 1$ mm³, 176 sagittal slices) along with a high-resolution whole-brain, cardiac-gated⁷ DTI sequence (TR_{eff} ~ 900 ms, TE = 92 ms, flip angle = 90° , FOV = 230×230 mm², voxel size = $1.983 \times 1.983 \times 2.000$ mm³, 72 interleaved slices with 60 gradient

directions and 10 non-diffusion weighted ($b = 0$) volumes, b value = 1500 s/mm²).

2.4. Data pre-processing

Pre-processing was completed through the Advanced Diffusion Pre-processing Pipeline developed in the Kurani AI & Neuroimaging Laboratory (Kurani, 2020) and processed via the Northwestern University Neuroimaging Data Archive (NUNDA) (Alpert et al., 2016). Steps included (1) denoising using principal components and creation of a b_0 reference image from the mean of the non-diffusion weighted scans; (2) skull stripping of the T1 structural image; (3) rigidly aligning a pseudo-T2 image (created by inverting the T1 image contrast) to the b_0 ; (4) nonlinear distortion correction; (5) eddy current correction and application of the b -vector file to yield rotated b -vectors; (6) concatenation of eddy current corrected parameters with the b_0 distortion field that was then applied to the diffusion scans; and (7) calculation of the diffusion tensor using the nonlinear weighted positive definite tensor-fitting algorithm from Camino (Cook et al., 2006). For data from three scans not processed in the above pipeline,⁸ data were processed using FSL 5.0.9 and SPM8. Specifically, for the diffusion-weighted data, steps included (1) creation of a brain mask using FSL's Brain Extraction Tool (bet2); (2) eddy current correction; and (3) calculation of diffusion tensors. For the T1 structural image, AC-PC alignment and skull stripping were completed. Subsequently, the processed diffusion-weighted data and the T1 structural image in native space were co-registered. The tensor files were used as inputs for Automated Fiber Quantification (AFQ).

2.5. Automated fiber quantification (AFQ)

AFQ, an open-source MATLAB-based tractography software that generates microstructural integrity metrics along white matter tracts, was then used for tractography and quantification of tract profiles. This included FA and MD of the CC major, CC minor, and bilateral AF, SLF, ILF, IFOF, and UF (Yeatman et al., 2012) (with use of MATLAB Version 2017a and SPM8). AFQ has been used to study white matter microstructural integrity in neurotypical adults (Angelopoulos et al., 2020; Zhou et al., 2018) as well as in a variety of brain disorders in adults and children, including dyslexia (Banfi et al., 2019), mild traumatic brain injury (Goodrich-Hunsaker et al., 2018) and aphasia (McKinnon et al., 2017, 2018; Zhang et al., 2018). Briefly, as detailed in Yeatman et al. (2012), AFQ completes deterministic tractography via the following steps:

a. Fiber tract identification via three steps: (1) "fiber tractography... using a deterministic streamlines tracking algorithm...seeded with a white matter mask defined as all voxels with FA greater than 0.3"; (2) waypoint region-of-interest (ROI)-based fiber tract segmentation, with waypoint ROIs drawn in MNI space and transformed into the individual's native space; and (3) fiber tract refinement based on comparison to a probabilistic fiber tract atlas with fiber tract probability maps transformed into the individual's native space.

b. Fiber tract cleaning for removal of fibers that deviate excessively from others in the same tract (i.e., more than 4 standard deviations above the mean fiber length or more than 5 standard deviations from the fiber tract core). Furthermore, additional manual tract cleaning was completed to eliminate implausible streamlines using QUENCH, a tool that allows for viewing and editing of tracts (Agrawal et al., 2011). This process was aided by visual comparison to tracts from healthy individuals in the freely-available AFQ browser (Yeatman et al., 2018). When delineated tracts were not neurobiologically plausible after automated cleaning and manual cleaning, these tracts were also eliminated from the analyses and treated as missing data (see Supplementary

⁵ For the 11 no-treatment natural history control participants, four never received treatment and seven served as treated participants following their natural history measurement period. One natural history control participant dropped out of the study before post-testing.

⁶ Data collection parameters were matched as closely as possible across scanners and sites to promote similar image quality. For participants at the Center for Translational Imaging, T1-weighted parameters differed as follows: TE = 2.94 ms.

⁷ Cardiac gating was employed using the pulse sensor to reduce the pulsatile motion which is more prevalent in elderly and stroke participants due to the increased amount of CSF. The effective TR (TR_{eff}) is based on the total time to collect a single direction which varied within and across subjects but was sufficiently long that T1 relaxation effects were not an issue (TR set to beat to beat interval of ~900 ms allowing 6 slices to be collected per heartbeat which required ~12 concatenations (heartbeats) to collect all 72 slices; effective TR = 12×0.9 s = 10.8 s).

⁸ P26 Timepoint 1 and 2 and C4 Timepoint 1 did not run in the above-mentioned pipeline due to temporary failure of supercomputer hardware.

Table 2

Linear regressions predicting WAB-R AQ for averaged white matter tract metrics at Threshold 2.

	Tract	N ^a	FA			-	MD		
			R ²	p ^b	q ^c		R ²	p ^b	q ^c
	CC major	27	0.012	0.582	–		0.033	0.365	–
	CC minor	33	0.049	0.213	–		0.005	0.689	–
Left	AF	10	0.563	0.012 [*]	0.074		0.559	0.013 [*]	0.078
	SLF	21	0.235	0.026 [*]	0.104		0.079	0.218	–
	ILF	29	0.422	<0.001 [*]	0.002 [*]		0.359	<0.001 [*]	0.007 [*]
	IFOF	14	0.189	0.12	–		0.265	0.060	–
	UF	14	0.001	0.91	–		<0.001	0.967	–
Right	AF	30	0.017	0.491	–		<0.001	0.963	–
	SLF	34	0.026	0.358	–		0.014	0.501	–
	ILF	34	0.001	0.844	–		0.002	0.826	–
	IFOF	34	0.002	0.782	–		0.005	0.679	–
	UF	33	0.01	0.571	–		0.001	0.856	–

^{*}statistically significant at p < .05.^{*}statistically significant at q < 0.05.

left hemisphere and corpus callosum forceps major and minor processed at Threshold 2 (lowered threshold); right hemisphere processed at Threshold 1 (standard AFQ threshold).

^a sample size based on delineated tracts for each of 12 tracts of interest.^b unadjusted p-values.^c q-values after FDR correction for 12 comparisons (depicted for items with significant p-values only); Note: none significant after controlling for lesion volume.

Table 2).

c. Fiber tract clipping to include the central portion of the tract between ROIs (i.e., fibers that extend beyond the defining ROIs are removed);

d. Fiber tract quantification through resampling fibers to 100 equidistant nodes along each tract. DTI scalars (i.e., MD and FA) for each node were calculated by taking a weighted average of diffusion properties for each fiber at that particular node (Yeatman et al., 2012). For the SLF, ILF, and IFOF, node 1 corresponds to the posterior part of the tract core and node 100 to the anterior part of the tract core. For the AF, node 1 corresponds to the frontal part of the tract core and node 100 corresponds to the temporal part of the tract core. For the UF, node 1 corresponds to the anterior temporal end of the tract core and node 100 corresponds to the frontal end of the tract core. The CC forceps major refers to the part of the corpus callosum connecting bi-hemispheric occipital regions and the CC forceps minor refers to the part of the corpus callosum connecting bi-hemispheric anterior frontal regions. For both CC tracts, node 1 corresponds to the end of the left-hemispheric part of the tract core and node 100 corresponds to the end of the right-hemispheric part of the tract core.

In addition to the procedures described above, two thresholds were evaluated to detect damaged tracts. For tract identification in AFQ, the default minimum tract length is 50 mm and individual streamlines are terminated when (a) FA < 0.2 and (b) the minimum angle between the last path segment and the next step direction is greater than 30° (Yeatman et al., 2012) (subsequently referred to as “Threshold 1”). These standard tractography parameters resulted in missing data in the left hemisphere for several of the participants. Thus, the minimum tract length was reduced to 20 mm and streamline termination criteria were reduced to FA < 0.1 and minimum angle greater than 35° in the left hemisphere to include a larger participant sample (subsequently referred to as “Threshold 2”). As mentioned in the introduction, these termination criteria vary across studies. The termination criteria selected here for both Threshold 1 and Threshold 2 have been used in previous work in people with aphasia (see McKinnon et al., 2018 for use of Threshold 2 parameters; see Zhang et al., 2018 for use of Threshold 1 FA termination criterion).

To evaluate the effect of lesion volume on brain-behavior correlations, lesion maps were manually drawn by a trained technician blinded to the behavioral data. Lesion maps were delineated from T1 structural data in native space using MRIcron software (Rorden & Brett, 2000). The

lesion tracing was based on the T1 image viewed in three orthogonal views simultaneously to improve accurate identification of the lesion. Frankly-damaged tissue was included while other structural abnormalities were excluded from the manual lesion identification. The lesion maps were then non-linearly warped to 1x1x1 mm³ MNI space (McGill’s MNI 2009c symmetric template) using the Advanced Diffusion Pre-processing Pipeline (Kurani, 2020) and processed via the Northwestern University Neuroimaging Data Archive (Alpert et al., 2016). Lesion volumes for all participants were then calculated using in-house MATLAB scripts (as described in (Meier et al., 2019b)).

2.6. Statistical analysis

For each tract, fractional anisotropy (FA) (ranging from 0 [diffuse] to 1 [coherent]) and mean diffusivity (MD) $((\lambda_1 + \lambda_2 + \lambda_3)/3)$; average of the three eigenvalues where lower values indicate lower water movement) were indexed by: (1) pointwise statistics of 100 equidistant nodes along the tract core (yielding 100 FA and MD values per tract) and (2) average values across all 100 nodes for each tract (yielding one FA and MD value per tract). R Version 3.6.2 was used for statistical analysis. Prior to independent samples t-tests, F-tests were completed to determine whether equal variances could be assumed.

For Question 1a, evaluating the relationship between tract delineation and language skills, independent samples t-tests were used to compare aphasia severity between participants with and without tracts delineated across the five left-hemisphere tracts at Threshold 1 and Threshold 2. For Questions 1b and 1c, regression models were used to determine whether microstructural integrity (i.e., average FA and MD values) across the 12 tracts of interest were related to WAB-R AQ and BNT-2 proportion correct at Timepoint 1. For prediction of BNT-2 proportion correct, given a bimodal distribution of this dataset via visual inspection, scores were binarized at a cutoff of 0.4 and logistic regression was completed. Pearson’s correlations were used to examine the relationship between pointwise microstructural integrity (i.e., FA and MD) and aphasia severity (i.e., WAB-R AQ) as well as pointwise microstructural integrity and naming impairment (i.e., BNT-2 proportion correct).

For Question 2, regression models were used to evaluate the relationship between average microstructural integrity (i.e., FA and MD) and treatment gains as measured by PMG (described in Section 2.2; for N = 30 participants who underwent treatment). Pearson’s correlations

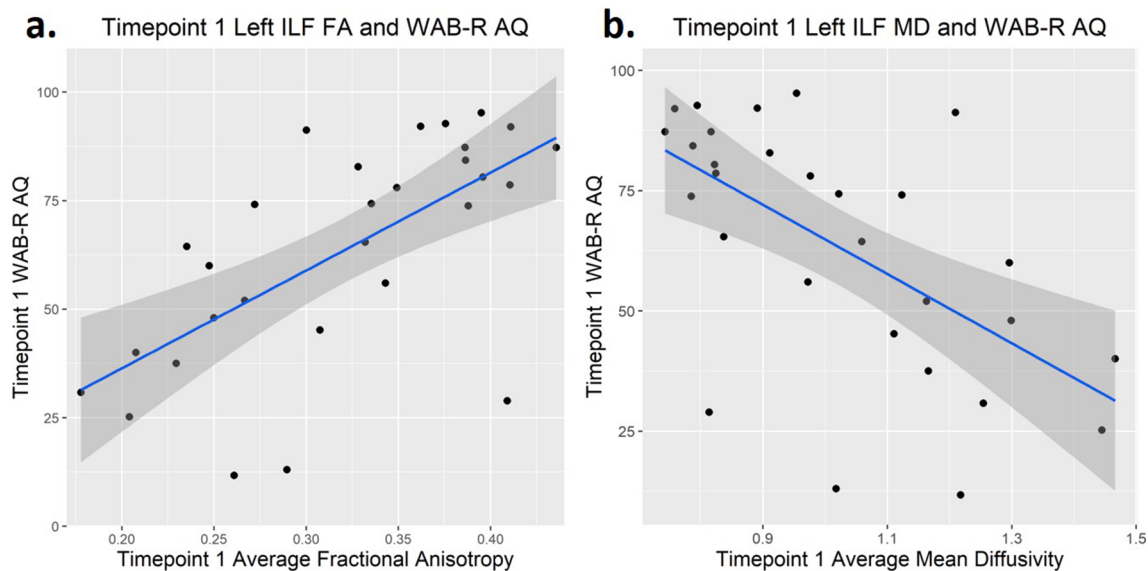


Fig. 1. Scatterplots of average diffusion measures of the left ILF and WAB-R AQ scores with regression lines and 95 % confidence intervals at Timepoint 1 for Threshold 2 for (a) FA and (b) MD ($n = 29$ with tracts delineated at Threshold 2).

were used for pointwise comparisons between DTI scalars and PMG. For Question 3, pre- to post-treatment changes in average and pointwise white matter microstructural integrity were evaluated with paired samples t-tests (for $N = 30$ participants who underwent treatment and separately for $N = 10$ no-treatment natural history control participants).

For all questions, correction for multiple comparisons was completed using the False Discovery Rate (FDR) procedure in R (Benjamini & Hochberg, 1995).⁹ Analyses with average measures were corrected for multiple comparisons using FDR correction applied for five comparisons for the question of tract delineation and 12 comparisons for each set of regression models relating neural data to language. Correlations with pointwise measures were corrected for multiple comparisons within each tract evaluated (i.e., 100 comparisons each). Results were considered significant below an alpha level of 0.05 (after FDR correction). Significant correlations and regression predictors were followed up with partial correlations accounting for lesion volume.

3. Results

Participants had a mean WAB-R AQ of 62.7 ($SD = 24.8$), mean BNT-2 score of 25/60 ($SD = 20$), mean pre-treatment baseline confrontation naming probe scores of 28.2 % ($SD = 21.8$), mean post-treatment confrontation naming probe scores of 52.9 % ($SD = 35.3$), and mean PMG of 0.43 ($SD = 0.36$) (see Supplementary Table 1 for individual evaluation and confrontation naming probe scores and Supplementary Fig. 1 for lesion overlays).

3.1. Question 1

3.1.1. Question 1a

At Threshold 1, left-hemisphere tracts could not be delineated for some participants due to tract damage (left AF 8/34, left SLF 20/34, left ILF 19/34, left IFOF 9/34, and left UF 9/34; see Supplementary Table 2 for complete list). Individuals with the left AF delineated showed significantly higher WAB-R AQ ($M = 84.5$, $SD = 10.0$) than those without the left AF delineated ($M = 56.0$, $SD = 24.3$, FDR -corrected $q <$

0.001). No other WAB-R AQ comparisons by left-hemisphere tract delineation were statistically significant at Threshold 1 (see Supplementary Table 4 and Supplementary Fig. 2). This result confirms the premise that individuals without delineated left-hemisphere tracts have more severe aphasia.

At Threshold 2, an increased number of left-hemisphere tracts could be delineated (left AF 10/34, left SLF 21/34, left ILF 29/34, left IFOF 14/34, left UF 14/34; see Supplementary Table 2 for complete list). T-tests comparing WAB-R AQ at Timepoint 1 for participants with and without tracts delineated showed significantly higher WAB-R AQ for those with tracts delineated vs not delineated for the left AF, IFOF, and UF (see Supplementary Table 4). No other WAB-R AQ comparisons by left-hemisphere tract delineation were statistically significant at Threshold 2.

3.1.2. Question 1b

3.1.2.1. Average measures. Regression models evaluating prediction of WAB-R AQ from average FA and MD across the 12 tracts of interest at Timepoint 1 showed no significant results after correction at Threshold 1 (see Supplementary Table 5). With the left-hemisphere tracts evaluated at Threshold 2, higher average FA of the left ILF was associated with higher WAB-R AQ (see Table 2 and Fig. 1a for full statistical results). Similar to FA, also for Threshold 2, lower average MD of the left ILF was significantly associated with higher WAB-R AQ (see Table 2 and Fig. 1b for full statistical results). When lesion volume was added as a covariate, no tract metrics were significant predictors.

3.1.2.2. Pointwise measures. At Threshold 1, Pearson's correlations showed that higher WAB-R AQ was associated with higher FA of the left SLF nodes 4–34 (FDR -corrected $q < 0.05$, r range = [0.534, 0.597], $n = 20$; see Figure panel 2b). These did not remain significant with a partial correlation accounting for lesion volume. At Threshold 2, Pearson's correlations showed that higher WAB-R AQ was associated with higher FA values of the left AF (nodes 1–68 and 76–82 with FDR -corrected $q < 0.05$, r range = [0.674, 0.869], $n = 10$), left SLF (nodes 5–33 and 82–95 with FDR -corrected $q < 0.05$, r range = [0.5, 0.592], $n = 21$), and left ILF (nodes 1–95 with FDR -corrected $q < 0.05$, r range = [0.376, 0.627], $n = 29$; see Figure panel 2c; none significant with partial correlation accounting for lesion volume).

With regard to pointwise MD measures, at Threshold 1, higher WAB-

⁹ For thoroughness, permutation-based multiple comparison correction (Nichols & Holmes, 2002; Yeatman et al., 2012) was also completed for analyses in Question 1 and Question 2 involving pointwise comparisons. The largely similar results are outlined in Supplementary Table 3.

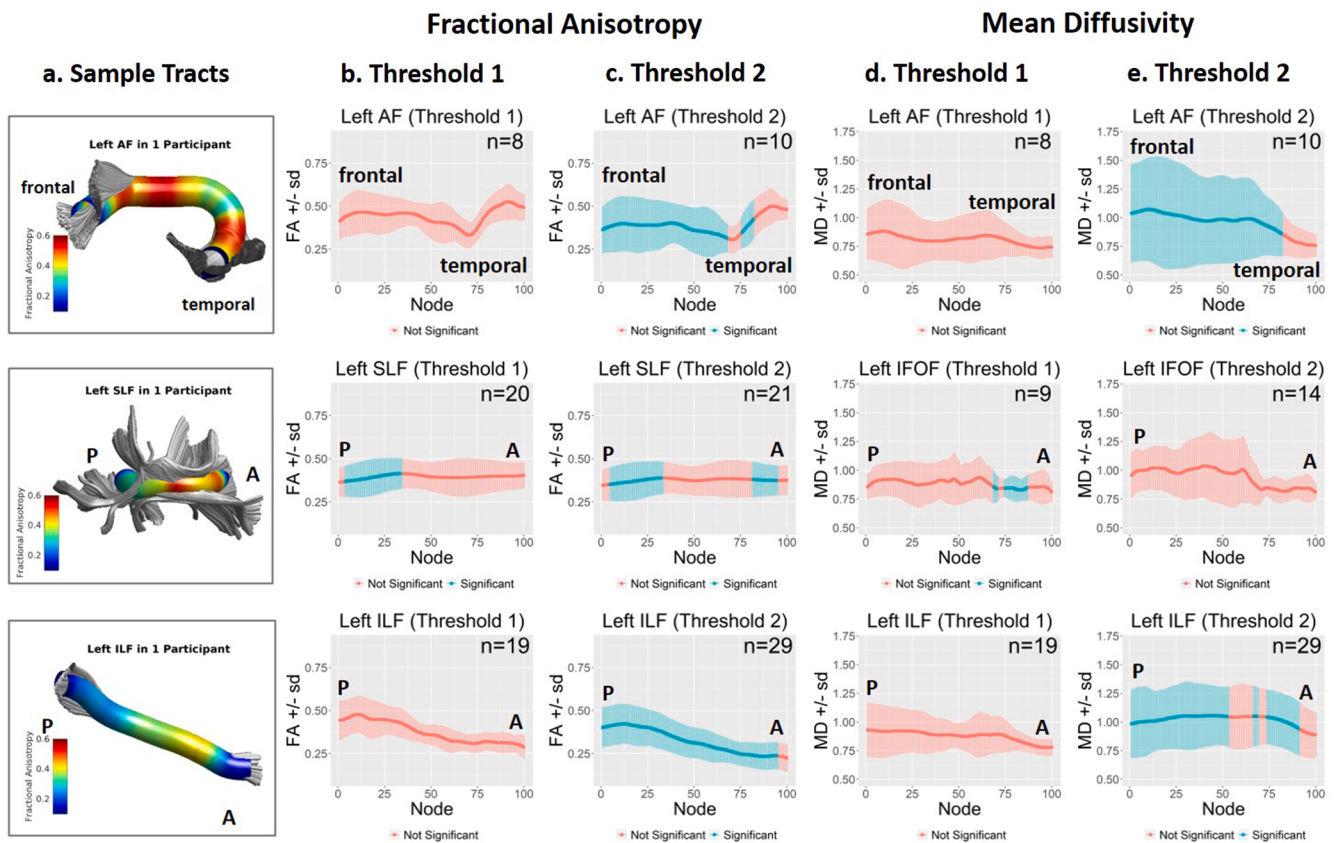


Fig. 2. (a) Sample tract rendering with along-tract fractional anisotropy at Threshold 2 for 1 participant; Pointwise along-tract measures for tracts with significant correlations across participants for (b) FA at Threshold 1, (c) FA at Threshold 2, (d) MD at Threshold 1, and (e) MD at Threshold 2. Nodes in teal are significantly correlated with WAB-R AQ, with positive correlations for FA (higher FA values associated with higher WAB-R AQ score) and negative correlations for MD (higher MD values associated with lower WAB-R AQ score). Error bars indicate $\pm$ standard deviation. Abbreviations: A (anterior), P (posterior).

R AQ was associated with lower MD of the left IFOF nodes 69–71 and 75–87 (*FDR-corrected* $q < 0.05$, r range = $[-0.958$,

$-0.813]$; see Figure panel 2d), which remained significant in a partial correlation accounting for lesion volume (*FDR-corrected* $q < 0.05$, r range = $[-0.986, -0.847]$). At Threshold 2, Pearson's correlations showed that higher WAB-R AQ was associated with lower MD values of the left AF (nodes 1–82 with *FDR-corrected* $q < 0.05$, r range = $[-0.804, -0.673]$, $n = 10$) and left ILF (nodes 1–53, 67–69, and 74–91 with *FDR-corrected* $q < 0.05$, r range = $[-0.591, -0.392]$, $n = 29$; see Figure panel 2e). None of these correlations were significant in partial correlations accounting for lesion volume.

3.1.3. Question 1c

3.1.3.1. Average measures. For average FA and MD, no tracts were significant predictors of binarized BNT score at either Threshold 1 or Threshold 2.

3.1.3.2. Pointwise measures. For Threshold 1, there were no significant correlations between pointwise measures and BNT proportion correct. At Threshold 2 in the left hemisphere, for pointwise FA across the 12 tracts of interest at Timepoint 1, higher BNT-2 scores were associated with higher FA values of the left ILF (nodes 1–25, 30–50 with *FDR-corrected* $q < 0.05$, r range = $[0.421, 0.542]$, $n = 29$; see Fig. 3a for pointwise plots). For pointwise MD across the 12 tracts of interest at Threshold 2, higher BNT-2 scores were associated with lower MD values of the left ILF (nodes 1–45 with *FDR-corrected* $q < 0.05$, r range = $[-0.59, -0.429]$, $n = 29$; see Fig. 3b for pointwise plots; not significant after partial correlation accounting for lesion volume).

3.2. Question 2

3.2.1. Average measures

For average FA and MD in the 12 tracts of interest at Timepoint 1, no tracts were significant predictors of treatment change for either Threshold 1 or Threshold 2.

3.2.2. Pointwise measures

For pointwise FA and MD at Threshold 1, there were no significant correlations between microstructural integrity metrics and treatment response. At Threshold 2, greater treatment response (i.e., PMG) was negatively associated with MD values of the left ILF at Timepoint 1 (nodes 1–43 with *FDR-corrected* $q < 0.05$, r range = $[-0.614, -0.459]$, $n = 29$; see Fig. 3c for pointwise plots).

3.3. Question 3

3.3.1. Average measures

For average FA and MD in the 12 tracts of interest, there were no significant differences from pre- to post-treatment at Threshold 1. At Threshold 2, paired samples t-tests showed that average FA value was significantly higher post-treatment ($M = 0.33$, $SD = 0.07$) vs pre-treatment for the left ILF ($M = 0.31$, $SD = 0.07$, *FDR-corrected* $q = 0.017$, $n = 23$; see Fig. 4a for boxplot). As discussed above, a smaller group of no-treatment natural history controls ($N = 10^{10}$) were included and paired samples t-tests showed no significant differences between

¹⁰ One no-treatment natural history control dropped out of the study before the post-treatment testing.

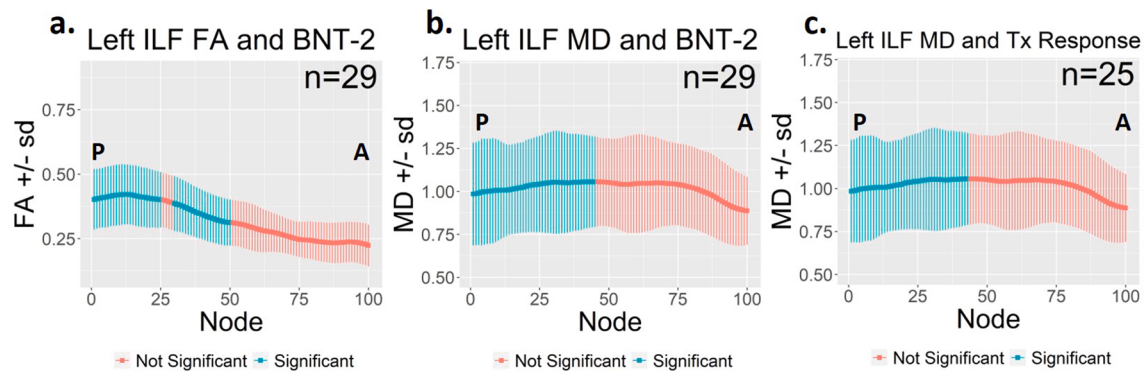


Fig. 3. Tract profiles for the left ILF including pointwise FA and MD across participants correlating with language scores at Threshold 2 for (a) FA along the left ILF and BNT-2 ($n = 29$ tracts delineated), (b) MD along the left ILF and BNT-2 ($n = 29$ tracts delineated), and (c) MD along the left ILF and treatment response (i.e., PMG) ($n = 25$ tracts delineated). Nodes in teal are significantly correlated with language scores at Threshold 2, with positive correlations for FA (higher FA values associated with higher language scores) and negative correlations for MD (higher MD values associated with lower language scores). Error bars indicate $\pm$ standard deviation. Abbreviations: A (anterior), P (posterior).

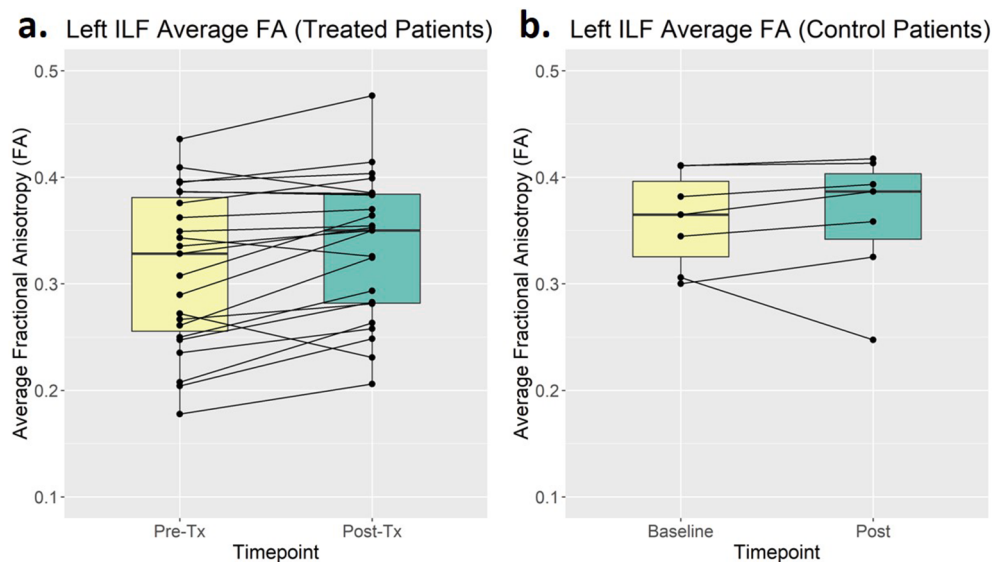


Fig. 4. Boxplots for pre- and post-treatment average FA at Threshold 2 with lines connecting paired datapoints in the (a) treatment group ($n = 25$ delineated at pre-treatment; $n = 25$ delineated at post-treatment; $n = 23$ available for paired statistical test) and (b) control group ($n = 8$ delineated at pre-treatment; $n = 7$ delineated at post-treatment; $n = 7$ available for paired statistical test).

Timepoint 1 and Timepoint 2 for averaged measures in any of the 12 tracts of interest at either Threshold 1 or Threshold 2 (see Fig. 4b for boxplot).

3.3.2. Pointwise measures

There were no significant changes from Timepoint 1 to Timepoint 2 for pointwise measures in the treated group at Threshold 1 or Threshold 2.

See Supplementary Table 6 for average FA and MD values at Threshold 1 and 2 across tracts and timepoints.

4. Discussion

This study evaluated microstructural and macrostructural integrity of white matter tracts and their relationship with language impairment and treatment response in individuals with chronic post-stroke aphasia. Along with average tract measures, the study also evaluated pointwise measures, which few previous studies in individuals with aphasia have done. Question 1 showed that individuals with fewer tracts delineated presented with more severe language impairment. With this finding as a

backdrop, there were several noteworthy results once the threshold of tract delineation was lowered with the goal of including data from tracts with more damage than allowed by the standard threshold.

The following were the main findings when considering the data evaluated at Threshold 2 (with a larger number of tracts delineated): First, average FA and MD in the left ILF were significant predictors of aphasia severity (i.e., WAB-R AQ). Additionally, pointwise FA of select tract nodes in the left AF, SLF, and ILF and pointwise MD for select nodes in the left AF and ILF were correlated with aphasia severity. For naming via the BNT-2, no average DTI scalars were significant predictors. Pointwise measures in select nodes of the left ILF were significantly correlated with naming (FA and MD) and treatment response (MD only). These results suggest that individuals with greater microstructural integrity of left-hemisphere tracts, and in particular the posterior left ILF, tend to have milder aphasia, greater naming skills, and more favorable response to treatment. Finally, participants in the treatment group, but not the no-treatment control group, showed a pre- to post-treatment increase in FA of the left ILF which suggests that typicality-based semantic feature treatment may have the capacity to alter the microstructural integrity of ventral stream tracts.

The first result, showing the difference in aphasia severity between participants with and without delineation of the left AF, aligns with previous reports that tract delineation of the left AF positively correlates with performance for global measures (Kim & Jang, 2013; Tak & Jang, 2014). Consequently, including FA or MD values from delineated tracts using typical tractography parameters may exclude individuals with more severe aphasia and more extensive left-hemisphere damage, making results less representative of the entire sample. Variability in tract delineation further complicates this matter (e.g., at Threshold 1, 9/34 tracts in the left UF were delineated versus 19/34 in the left ILF). From a statistical standpoint, this vast difference presents challenges to building comprehensive predictive models that incorporate microstructural integrity metrics from multiple tracts. In order to include data from more damaged tracts, and subsequently more participants, lowered tractography thresholds were used in the left hemisphere. However, further work needs to be completed to determine optimal thresholds and this caveat should be considered in interpreting subsequent results.

Findings that both average and pointwise measures for the left AF, SLF, and ILF relate to WAB-R AQ align with previous studies correlating language performance to integrity of dorsal (e.g., Ivanova et al., 2016; Keser, Meier, Stockbridge, & Hillis, 2020a; Kim & Jang, 2013; J.-K. Lee et al., 2021; S. Lee et al., 2020; McKinnon et al., 2018; Yang et al., 2017; Zhang et al., 2021) and ventral tracts (e.g., Ivanova et al., 2016; McKinnon et al., 2018; Meier et al., 2019b; Xing et al., 2017; Zhang et al., 2018). These results also align with the dual-stream theory (Hickok & Poeppel, 2004, 2007), given that both dorsal and ventral stream integrity are important for a global measure of language function that incorporates many different linguistic skills. As noted above, interpretation of findings should be in consideration of the small number of tracts delineated for the left AF, even at Threshold 2. In terms of treatment response, the finding that Timepoint 1 MD values of the left ILF negatively correlate with treatment response suggests that pre-treatment microstructural integrity of this ventral stream tract may be able to provide prognostic information; namely, lower MD values (i.e., suggestive of lower diffusion and thus higher microstructural integrity) predict greater response to semantically-based anomia treatment.

Overall, evidence in the current study for relationships between left-hemisphere microstructural integrity and language performance verifies and also extends previous findings in individuals with aphasia by providing fine-grained information about integrity of specific tract nodes that relate to behavior. Across analyses, the posterior left ILF emerged as a key predictor in this dataset. The middle and superior temporal gyri and structural connectivity provided by the ILF have previously been implicated in semantic processing (Turken & Dronkers, 2011). It may be the case that structural connections from the posterior ILF are a particularly important part of the ventral stream for support of lexical-semantic processing (e.g., due to occipito-temporal structural connectivity needed for integration of visual and semantic information (Panesar et al., 2018)). Nonetheless, information provided about the role of the ILF from this analysis may be incomplete given that the current approach does not account for anatomically-based subdivisions of the tract, such as dorsal and ventral sub-fascicles (Panesar et al., 2018) or multiple branches with differing gray matter connections (Mandonnet et al., 2018). Review of these anatomical subdivisions in future work could provide additional insight.

Pre- to post-treatment increases in average FA in the left ILF suggest that the left ILF may undergo changes in response to semantically-based treatment. While studies evaluating microstructural changes in individuals with chronic post-stroke aphasia have focused on the AF, there is some evidence to suggest microstructural plasticity of the left ILF (McKinnon et al., 2017) and left-hemisphere gray matter ventral stream regions (Chang et al., 2021). In the current study, a connection between semantically-based treatment and ILF integrity changes may indicate that a lexical-semantic treatment targeting skills presumably supported by the ILF relates to changes in the ILF. Nonetheless, these results should be interpreted with caution given the smaller sized control group.

Furthermore, the larger amount of data available for the left ILF compared to other left-hemisphere tracts also may have contributed to the disparities in statistical significance. However, it should be noted that properties of right-hemisphere tracts, for which there were little missing data, were consistently not significant predictors of language performance. This is in contrast to a few other studies in the literature which have found evidence of a relationship between integrity of right-hemisphere white matter microstructural integrity and language skills (Forkel et al., 2014; Hartwigsen et al., 2020; Osa García et al., 2020). This discrepancy could be due to methodological differences between studies or differences in sample characteristics.

Across analyses, associations did not remain after accounting for lesion volume. This confirms the strong relationship of lesion size with aphasia severity and treatment prognosis (Meier et al., 2019a). Even so, as others have pointed out, information about tract integrity and its relationship to behavioral measures potentially provides additional information beyond just overall predictive power, including for development of treatments that target particular areas of the brain (Jang et al., 2017; McKinnon et al., 2017). In the future, treatments could be selected to improve structural and functional integrity of a particular brain region. For example, a semantically-based treatment such as semantic feature analysis might be used to target improved lexical-semantic access through improvement of underlying left ILF structure and function.

Several limitations of this study should be mentioned. First, DTI presents inherent limitations. One critical problem with a DTI model for the purposes of tractography is the inability to distinguish crossing fibers. The method used in the current study can only predict a single fiber population per voxel, and thus errors may occur when there are crossing fibers within a voxel (Jeurissen et al., 2019), particularly if there is a mix of damaged and undamaged fibers within a voxel. Researchers have started to address this issue in recent literature with more sophisticated data acquisition and processing techniques such as multi-shell diffusion models (Koniczny et al., 2021). Furthermore, FA and MD are indirect measures from which one infers microstructural properties. In particular, FA is susceptible to error due to the presence of edema in injured tissue. Previous work has also pointed out that considering higher FA and lower MD as indicative of better microstructural integrity may be overly simplistic, given that these measures vary within and between healthy tracts (Soares et al., 2013). In future studies, comparison to an age-matched neurotypical control group could help determine which tracts' metrics differ from typical values. This will be particularly important to confirm whether tract and tract segment metrics that correlate with language performance tend to have higher microstructural abnormality.

Next, the current investigation did not use lesion masks to remove lesioned voxels from the analysis. While removal of lesioned voxels from fMRI data is standard, removal of lesioned voxels from DTI data has varied in the literature. One methodological investigation found that FA values gathered from tract-based spatial statistics (TBSS) for the bilateral corticospinal tract and SLF were similar when comparing the procedure done with and without lesion masking, concluding that performing TBSS without lesion masks may be sufficient (Koyama et al., 2019). While not specifically examining use of lesion masks for tractography, this result suggests that methods designed to gather FA values may inherently filter out lesioned voxels, although this requires further investigation. This study was also limited by small sample size. This is an ongoing challenge with clinical populations and should continue to be addressed with multi-site studies and data sharing among research groups.

Last, some of the research questions yielded statistically significant results in pointwise measures only, some in averaged measures only, and some in both. It is possible that average measures (indexing overall tract integrity) are more informative in relation to evaluating neural changes and that pointwise measures (indexing integrity at precise location on the tract) may be more informative for relating neural structure to behavioral performance. Future work should examine these issues.

5. Conclusions

Overall, this study showed that left-hemisphere white matter microstructural integrity metrics from diffusion tensor imaging relate to aphasia severity, naming, and treatment response in individuals with chronic post-stroke aphasia both for averaged tract measures and in pointwise measures in nodes along tracts. The study also found preliminary evidence of improved microstructural integrity of the left ILF following anomia treatment. These results suggest that white matter microstructural integrity of left-hemisphere dorsal and ventral stream tracts should be considered in the future as potential features for models predicting aphasia severity and treatment response post-stroke, along with other previously studied measures of white matter function such as white matter hyperintensities (e.g. Varkanitsa et al., 2020; Wilmskoetter et al., 2019).

CRediT authorship contribution statement

Emily J. Braun: Data curation, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Anne Billot:** Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Erin L. Meier:** Conceptualization, Data curation, Investigation, Writing – review & editing. **Yue Pan:** Data curation, Software, Investigation, Writing – review & editing. **Todd B. Parrish:** Data curation, Methodology, Software, Writing – review & editing. **Ajay S. Kurani:** Data curation, Methodology, Software, Writing – review & editing. **Swathi Kiran:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

Swathi Kiran is a scientific consultant for Constant Therapy Health with no overlap between this role and the submitted investigation. The authors have no other competing interests to declare.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by NIH/NIDCD P50DC012283 and NIH/NIDCD T32DC013017. The authors acknowledge our research participants who generously gave their time and energy during evaluation and treatment for this study. The authors also acknowledge the members of the Aphasia Research Laboratory at Boston University for many helpful discussions and Jeffrey Johnson and Natalie Gilmore for data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bandl.2022.105163>.

References

- Agrawal, S., Akers, D., & Sherbondy, T. (2011). *QUENCH*. <https://web.stanford.edu/group/vista/cgi-bin/wiki/index.php/QUENCH>.
- Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2008). Diffusion Tensor Imaging of the Brain. *Neurotherapeutics*, 4(3), 316–329.
- Alpert, K., Kogan, A., Parrish, T., Marcus, D., & Wang, L. (2016). The Northwestern University Neuroimaging Data Archive (NUNDA). *NeuroImage*, 124, 1131–1136. <https://doi.org/10.1016/j.neuroimage.2015.05.060>
- Angelopoulou, G., Meier, E. L., Kasselimis, D., Pan, Y., Tsolakopoulos, D., Velonakis, G., Karavasilis, E., Kelekis, N. L., Goutsos, D., Potagas, C., & Kiran, S. (2020). Investigating Gray and White Matter Structural Substrates of Sex Differences in the Narrative Abilities of Healthy Adults. *Frontiers in Neuroscience*, 13, 1424. <https://doi.org/10.3389/fnins.2019.01424>
- Banfi, C., Koschutnig, K., Moll, K., Schulte-Körne, G., Fink, A., & Landerl, K. (2019). White matter alterations and tract lateralization in children with dyslexia and isolated spelling deficits. *Human Brain Mapping*, 40(3), 765–776. <https://doi.org/10.1002/hbm.24410>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, 19(12), 2767–2796. <https://doi.org/10.1093/cercor/bhp055>
- Blom-Smink, M., Verly, M., Spielmann, K., Smits, M., Ribbers, G. M., & van de Sandt-Koenderman, M. W. M. E. (2020). Change in Right Inferior Longitudinal Fasciculus Integrity Is Associated With Naming Recovery in Subacute Poststroke Aphasia. *Neurorehabilitation and Neural Repair*, 34(9), 784–794. <https://doi.org/10.1177/1545968320940982>
- Bonilha, L., Gleichgerricht, E., Nesland, T., Rorden, C., & Fridriksson, J. (2016). Success of Anomia Treatment in Aphasia Is Associated With Preserved Architecture of Global and Left Temporal Lobe Structural Networks. *Neurorehabilitation and Neural Repair*, 30(3), 266–279. <https://doi.org/10.1177/1545968315593808>
- Campbell, J. S. W., & Pike, G. B. (2014). Potential and limitations of diffusion MRI tractography for the study of language. *Brain and Language*, 131, 65–73. <https://doi.org/10.1016/j.bandl.2013.06.007>
- Chang, A. J., Wilmskoetter, J., Fridriksson, J., McKinnon, E. T., Johnson, L. P., Basilakos, A., Jensen, J. H., Rorden, C., & Bonilha, L. (2021). Cortical microstructural changes associated with treated aphasia recovery. *Annals of Clinical and Translational Neurology*, 8(9), 1884–1894. <https://doi.org/10.1002/actn.3.51445>
- Cook, P. A., Bai, Y., Nedjati-Gilani, S., Seunarine, K. K., Hall, G., Parker, G. J., & Alexander, D. C. (2006). Camino: Open-Source Diffusion-MRI Reconstruction and Processing. *14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, 2759.
- Dresang, H. C., Hula, W. D., Yeh, F.-C., Warren, T., & Dickey, M. W. (2021). White-matter neuroanatomical predictors of aphasic verb retrieval. *Brain Connectivity*, 11(4), 319–330. <https://doi.org/10.1089/brain.2020.0921>
- Flowers, H. L., Skoretz, S. A., Silver, F. L., Rochon, E., Fang, J., Flamand-Roze, C., & Martino, R. (2016). Poststroke Aphasia Frequency, Recovery, and Outcomes: A Systematic Review and Meta-Analysis. *Archives of Physical Medicine and Rehabilitation*, 97(12), 2188–2201.e8. <https://doi.org/10.1016/j.apmr.2016.03.006>
- 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(7), 2027–2039. <https://doi.org/10.1093/brain/awu113>
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84. [https://doi.org/10.1016/S1364-6613\(00\)01839-8](https://doi.org/10.1016/S1364-6613(00)01839-8)
- Friederici, A. D. (2011). The Brain Basis of Language Processing: From Structure to Function. *Physiological Reviews*, 91(4), 1357–1392. <https://doi.org/10.1152/physrev.00006.2011>
- Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–268. <https://doi.org/10.1016/j.tics.2012.04.001>
- Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (2007). Role of the Corpus Callosum in Speech Comprehension: Interfacing Syntax and Prosody. *Neuron*, 53(1), 135–145. <https://doi.org/10.1016/j.neuron.2006.11.020>
- 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 and Language*, 150, 117–128. <https://doi.org/10.1016/j.bandl.2015.09.001>
- Geva, S., Correia, M., & Warburton, E. A. (2011). Diffusion tensor imaging in the study of language and aphasia. *Aphasiology*, 25(5), 543–558. <https://doi.org/10.1080/02687038.2010.534803>
- Gilmore, N., Meier, E. L., Johnson, J. P., & Kiran, S. (2019). Nonlinguistic Cognitive Factors Predict Treatment-Induced Recovery in Chronic Poststroke Aphasia. *Archives of Physical Medicine and Rehabilitation*, 100(7), 1251–1258. <https://doi.org/10.1016/j.apmr.2018.12.024>
- Gilmore, N., Meier, E. L., Johnson, J. P., & Kiran, S. (2020). Typicality-based semantic treatment for anomia results in multiple levels of generalisation. *Neuropsychological Rehabilitation*, 30(5), 802–828. <https://doi.org/10.1080/09602011.2018.1499533>
- Goodglass, H., Kaplan, E., & Weintraub, S. (2001). *Boston Naming Test-Second Edition*. Pro-Ed.
- Goodrich-Hunsaker, N. J., Abildskov, T. J., Black, G., Bigler, E. D., Cohen, D. M., Mihalov, L. K., Bangert, B. A., Taylor, H. G., & Yeates, K. O. (2018). Age- and sex-related effects in children with mild traumatic brain injury on diffusion magnetic resonance imaging properties: A comparison of voxelwise and tractography methods. *Journal of Neuroscience Research*, 96(4), 626–641. <https://doi.org/10.1002/jnr.24142>
- Han, Z., Ma, Y., Gong, G., He, Y., Caramazza, A., & Bi, Y. (2013). White matter structural connectivity underlying semantic processing: Evidence from brain damaged patients. *Brain*, 136(10), 2952–2965. <https://doi.org/10.1093/brain/awt205>
- Han, Z., Ma, Y., Gong, G., Huang, R., Song, L., & Bi, Y. (2016). White matter pathway supporting phonological encoding in speech production: A multi-modal imaging study of brain damage patients. *Brain Structure and Function*, 221(1), 577–589. <https://doi.org/10.1007/s00429-014-0926-2>
- Hartwigsen, G., Stockert, A., Charpentier, L., Wawrzyniak, M., Klingbeil, J., Wrede, K., Obrig, H., & Saur, D. (2020). Short-term modulation of the lesioned language network. *eLife*, 9, Article e54277. <https://doi.org/10.7554/eLife.54277>

- 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. <https://doi.org/10.1016/j.cortex.2015.03.004>
- Harvey, D. Y., Wei, T., Ellmore, T. M., Hamilton, A. C., & Schnur, T. T. (2013). Neuropsychological evidence for the functional role of the uncinate fasciculus in semantic control. *Neuropsychologia*, 51(5), 789–801. <https://doi.org/10.1016/j.neuropsychologia.2013.01.028>
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1–2), 67–99. <https://doi.org/10.1016/j.cognition.2003.10.011>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402. <https://doi.org/10.1038/nrn2113>
- Hosomi, A., Nagakane, Y., Yamada, K., Kuriyama, N., Mizuno, T., Nishimura, T., & Nakagawa, M. (2009). Assessment of arcuate fasciculus with diffusion-tensor tractography may predict the prognosis of aphasia in patients with left middle cerebral artery infarcts. *Neuroradiology*, 51(9), 549–555. <https://doi.org/10.1007/s00234-009-0534-7>
- 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*, 1–13. <https://doi.org/10.1093/brain/awaa193>
- Ivanova, M. V., Isaev, D. Y., Dragoy, O. V., Akina, Y. S., Petrushevskiy, 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*, 85, 165–181. <https://doi.org/10.1016/j.cortex.2016.04.019>
- Jang, S. H., Cho, I. T., & Lim, J. W. (2017). Recovery of aphasia and change of injured arcuate fasciculus in the dominant hemisphere in stroke patients. *NeuroRehabilitation*, 41(4), 759–764. <https://doi.org/10.3233/NRE-172167>
- Jeurissen, B., Descoteaux, M., Mori, S., & Leemans, A. (2019). Diffusion MRI fiber tractography of the brain. *NMR in Biomedicine*, 32(4), Article e3785. <https://doi.org/10.1002/nbm.3785>
- Johnson, J. P., Meier, E. L., Pan, Y., & Kiran, S. (2019). Treatment-related changes in neural activation vary according to treatment response and extent of spared tissue in patients with chronic aphasia. *Cortex*, 121, 147–168. <https://doi.org/10.1016/j.cortex.2019.08.016>
- 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 and Language*, 207, Article 104809. <https://doi.org/10.1016/j.bandl.2020.104809>
- Kertesz, A. (2007). *Western Aphasia Battery—Revised*. Pearson.
- Keser, Z., Meier, E. L., Stockbridge, M. D., & Hillis, A. E. (2020a). The role of microstructural integrity of major language pathways in narrative speech in the first year after stroke. *Journal of Stroke and Cerebrovascular Diseases*, 29(9), Article 105078. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105078>
- Keser, Z., Sebastian, R., Hasan, K. M., & Hillis, A. E. (2020b). Right Hemispheric Homologous Language Pathways Negatively Predicts Poststroke Naming Recovery. *Stroke*, 51(3), 1002–1005. <https://doi.org/10.1161/STROKEAHA.119.028293>
- Kim, S. H., & Jang, S. H. (2013). Prediction of Aphasia Outcome Using Diffusion Tensor Tractography for Arcuate Fasciculus in Stroke. *American Journal of Neuroradiology*, 34(4), 785–790. <https://doi.org/10.3174/ajnr.A3259>
- Kiran, S., & Thompson, C. K. (2003). The Role of Semantic Complexity in Treatment of Naming Deficits: Training Semantic Categories in Fluent Aphasia by Controlling Exemplar Typicality. *Journal of Speech, Language, and Hearing Research*, 46(4), 773–787.
- Konieczny, M. J., Dewenter, A., ter Telgte, A., Gesierich, B., Wiegertjes, K., Finsterwalder, S., Kopczak, A., Hübner, M., Malik, R., Tuladhar, A. M., Marques, J. P., Norris, D. G., Koch, A., Dietrich, O., Ewers, M., Schmidt, R., de Leeuw, F.-E., & Duering, M. (2021). Multi-shell Diffusion MRI Models for White Matter Characterization in Cerebral Small Vessel Disease. *Neurology*, 96(5), Article e698. <https://doi.org/10.1212/WNL.00000000000011213>
- Koyama, T., Uchiyama, Y., & Domen, K. (2019). Comparison of Fractional Anisotropy from Tract-Based Spatial Statistics with and without Lesion Masking in Patients with Intracerebral Hemorrhage: A Technical Note. *Journal of Stroke and Cerebrovascular Diseases*, 28(11), Article 104376. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.104376>
- Kurani, A. S. (2020). April 30). *Advanced Diffusion Preprocessing Pipeline*. <https://www.kuranilab.fsm.northwestern.edu/software/adpp>
- Lambon Ralph, M. A., Snell, C., Fillingham, J. K., Conroy, P., & Sage, K. (2010). Predicting the outcome of anomia therapy for people with aphasia post CVA: Both language and cognitive status are key predictors. *Neuropsychological Rehabilitation*, 20(2), 289–305. <https://doi.org/10.1080/09602010903237875>
- Lätt, J., Nilsson, M., Wirestam, R., Ståhlberg, F., Karlsson, N., Johansson, M., Sundgren, P. C., & van Westen, D. (2013). Regional values of diffusional kurtosis estimates in the healthy brain. *Journal of Magnetic Resonance Imaging*, 37(3), 610–618. <https://doi.org/10.1002/jmri.23857>
- Lee, J.-K., Ko, M.-H., Park, S.-H., & Kim, G.-W. (2021). Prediction of Aphasia Severity in Patients with Stroke Using Diffusion Tensor Imaging. *Brain Sciences*, 11(3), 304. <https://doi.org/10.3390/brainsci11030304>
- Lee, S., Na, Y., Tae, W.-S., & Pyun, S.-B. (2020). Clinical and neuroimaging factors associated with aphasia severity in stroke patients: Diffusion tensor imaging study. *Scientific Reports*, 10(1), 12874. <https://doi.org/10.1038/s41598-020-69741-1>
- Mandonnet, E., Sarubbo, S., & Petit, L. (2018). The nomenclature of human white matter association pathways: Proposal for a systematic taxonomic anatomical classification. *Frontiers in Neuroanatomy*, 12, 94. <https://doi.org/10.3389/fnana.2018.00094>
- 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(8), 2251–2256. <https://doi.org/10.1161/STROKEAHA.110.606103>
- McKinnon, E. T., Fridriksson, J., Basilakos, A., Hickok, G., Hillis, A. E., Spampinato, M. V., Gleichgerricht, E., Rorden, C., Jensen, J. H., Helpen, J. A., & Bonilha, L. (2018). Types of naming errors in chronic post-stroke aphasia are dissociated by dual stream axonal loss. *Scientific Reports*, 8(1), 14352. <https://doi.org/10.1038/s41598-018-32457-4>
- McKinnon, E. T., Fridriksson, J., Glenn, G. R., Jensen, J. H., Helpen, J. A., Basilakos, A., Rorden, C., Shih, A. Y., Spampinato, M. V., & Bonilha, L. (2017). Structural plasticity of the ventral stream and aphasia recovery: Structural Plasticity and Aphasia Recovery. *Annals of Neurology*, 82(1), 147–151. <https://doi.org/10.1002/ana.24983>
- Meier, E. L., Johnson, J. P., & Kiran, S. (2018). Left frontotemporal effective connectivity during semantic feature judgments in patients with chronic aphasia and age-matched healthy controls. *Cortex*, 108, 173–192. <https://doi.org/10.1016/j.cortex.2018.08.006>
- Meier, E. L., Johnson, J. P., Pan, Y., & Kiran, S. (2019a). A lesion and connectivity-based hierarchical model of chronic aphasia recovery dissociates patients and healthy controls. *NeuroImage: Clinical*, 23, Article 101919. <https://doi.org/10.1016/j.nicl.2019.101919>
- Meier, E. L., Johnson, J. P., Pan, Y., & Kiran, S. (2019b). The utility of lesion classification in predicting language and treatment outcomes in chronic stroke-induced aphasia. *Brain Imaging and Behavior*, 13(6), 1510–1525. <https://doi.org/10.1007/s11682-019-00118-3>
- Mori, S. (2009). *Introduction to diffusion tensor imaging (Repr)*. Elsevier.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, 15(1), 1–25. <https://doi.org/10.1002/hbm.1058>
- Osa García, A., Brambati, S. M., Brisebois, A., Désilets-Barnabé, M., Houzé, B., Bedetti, C., Rochon, E., Leonard, C., Desautels, A., & Marcotte, K. (2020). Predicting Early Post-stroke Aphasia Outcome From Initial Aphasia Severity. *Frontiers in Neurology*, 11, 120. <https://doi.org/10.3389/fneur.2020.00120>
- Panesar, S. S., Yeh, F.-C., Jacquesson, T., Hula, W., & Fernandez-Miranda, J. C. (2018). A quantitative tractography study into the connectivity, segmentation and laterality of the human inferior longitudinal fasciculus. *Frontiers in Neuroanatomy*, 12, 47. <https://doi.org/10.3389/fnana.2018.00047>
- Pani, E., Zheng, X., Wang, J., Norton, A., & Schlaug, G. (2016). Right hemisphere structures predict poststroke speech fluency. *Neurology*, 86(17), 1574–1581. <https://doi.org/10.1212/WNL.0000000000002613>
- Papoutsis, M., Stamatakis, E. A., Griffiths, J., Marslen-Wilson, W. D., & Tyler, L. K. (2011). Is left fronto-temporal connectivity essential for syntax? Effective connectivity, tractography and performance in left-hemisphere damaged patients. *NeuroImage*, 58(2), 656–664. <https://doi.org/10.1016/j.neuroimage.2011.06.036>
- Pierpaoli, C., Jezzard, P., Basser, P. J., Barnett, A., & Di Chiro, G. (1996). Diffusion tensor MR imaging of the human brain. *Radiology*, 201(3), 637–648. <https://doi.org/10.1148/radiology.201.3.8939209>
- Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, 1191(1), 62–88. <https://doi.org/10.1111/j.1749-6632.2010.05444.x>
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, 62(2), 816–847. <https://doi.org/10.1016/j.neuroimage.2012.04.062>
- Ramsey, L. E., Siegel, J. S., Lang, C. E., Strube, M., Shulman, G. L., & Corbetta, M. (2017). Behavioural clusters and predictors of performance during recovery from stroke. *Nature Human Behaviour*, 1(3), 0038. <https://doi.org/10.1038/s41562-016-0038>
- Rolheiser, T., Stamatakis, E. A., & Tyler, L. K. (2011). Dynamic Processing in the Human Language System: Synergy between the Arcuate Fascicle and Extreme Capsule. *Journal of Neuroscience*, 31(47), 16949–16957. <https://doi.org/10.1523/JNEUROSCI.2725-11.2011>
- Rorden, C., & Brett, M. (2000). Stereotaxic Display of Brain Lesions. *Behavioural Neurology*, 12(4), 191–200. <https://doi.org/10.1155/2000/421719>
- Rosso, C., Vargas, P., Valabregue, R., Arbizu, C., Henry-Amar, F., Leger, A., Lehericy, S., & Samson, Y. (2015). Aphasia severity in chronic stroke patients: A combined disconnection in the dorsal and ventral language pathways. *Neurorehabilitation and Neural Repair*, 29(3), 287–295. <https://doi.org/10.1177/1545968314543926>
- Saur, D., Kreher, B. W., Schnell, S., Kummerer, 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. *Proceedings of the National Academy of Sciences*, 105(46), 18035–18040. <https://doi.org/10.1073/pnas.0805234105>
- Schlaug, G., Marchina, S., & Norton, A. (2009). Evidence for Plasticity in White-Matter Tracts of Patients with Chronic Broca's Aphasia Undergoing Intense Intonation-based Speech Therapy. *Annals of the New York Academy of Sciences*, 1169(1), 385–394. <https://doi.org/10.1111/j.1749-6632.2009.04587.x>
- Soares, J. M., Marques, P., Alves, V., & Sousa, N. (2013). A hitchhiker's guide to diffusion tensor imaging. *Frontiers in Neuroscience*, 7. <https://doi.org/10.3389/fnins.2013.00031>
- Tae, W.-S., Ham, B.-J., Pyun, S.-B., Kang, S.-H., & Kim, B.-J. (2018). Current Clinical Applications of Diffusion-Tensor Imaging in Neurological Disorders. *Journal of Clinical Neurology*, 14(2), 129. <https://doi.org/10.3988/jcn.2018.14.2.129>
- Tak, H. J., & Jang, S. H. (2014). Relation between aphasia and arcuate fasciculus in chronic stroke patients. *BMC Neurology*, 14(1), 46. <https://doi.org/10.1186/1471-2377-14-46>
- Turken, A. U., & Dronkers, N. F. (2011). The Neural Architecture of the Language Comprehension Network: Converging Evidence from Lesion and Connectivity Analyses. *Frontiers in Systems Neuroscience*, 5, 1. <https://doi.org/10.3389/fnsys.2011.00001>

- van Hees, S., McMahon, K., Angwin, A., de Zubicaray, G., Read, S., & Copland, D. A. (2014). Changes in White Matter Connectivity Following Therapy for Anomia Post stroke. *Neurorehabilitation and Neural Repair*, 28(4), 325–334. <https://doi.org/10.1177/1545968313508654>
- Varkanitsa, M., Peñaloza, C., Charidimou, A., Caplan, D., & Kiran, S. (2020). White Matter Hyperintensities Predict Response to Language Treatment in Poststroke Aphasia. *Neurorehabilitation and Neural Repair*, 34(10), 945–953. <https://doi.org/10.1177/1545968320952809>
- Virani, S. S., Alonso, A., Aparicio, H. J., Benjamin, E. J., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Cheng, S., Delling, F. N., Elkind, M. S. V., Evenson, K. R., Ferguson, J. F., Gupta, D. K., Khan, S. S., Kissela, B. M., Knutson, K. L., Lee, C. D., Lewis, T. T., ... On behalf of the American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2021). Heart Disease and Stroke Statistics—2021 Update: A Report From the American Heart Association. *Circulation*, 143(8). <https://doi.org/10.1161/CIR.0000000000000950>
- Wan, C. Y., Zheng, X., Marchina, S., Norton, A., & Schlaug, G. (2014). Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain and Language*, 136, 1–7. <https://doi.org/10.1016/j.bandl.2014.03.011>
- Wang, J., Marchina, S., Norton, A. C., Wan, C. Y., & Schlaug, G. (2013). Predicting speech fluency and naming abilities in aphasic patients. *Frontiers in Human Neuroscience*, 7, 831. <https://doi.org/10.3389/fnhum.2013.00831>
- Wilmskoetter, J., Marebwa, B., Basilakos, A., Fridriksson, J., Rorden, C., Stark, B. C., Johnson, L., Hickok, G., Hillis, A. E., & Bonilha, L. (2019). Long-range fibre damage in small vessel brain disease affects aphasia severity. *Brain: A Journal of Neurology*, 142(10), 3190–3201. <https://doi.org/10.1093/brain/awz251>
- Xing, S., Lacey, E. H., Skipper-Kallal, L. M., Zeng, J., & Turkeltaub, P. E. (2017). White Matter Correlates of Auditory Comprehension Outcomes in Chronic Post-Stroke Aphasia. *Frontiers in Neurology*, 8, 54. <https://doi.org/10.3389/fneur.2017.00054>
- Yang, M., Li, Y., Li, J., Yao, D., Liao, W., & Chen, H. (2017). Beyond the Arcuate Fasciculus: Damage to Ventral and Dorsal Language Pathways in Aphasia. *Brain Topography*, 30(2), 249–256. <https://doi.org/10.1007/s10548-016-0503-5>
- Yeatman, J. D., Dougherty, R. F., Myall, N. J., Wandell, B. A., & Feldman, H. M. (2012). Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification. *PLoS ONE*, 7(11), Article e49790. <https://doi.org/10.1371/journal.pone.0049790>
- Yeatman, J. D., Richie-Halford, A., Smith, J. K., Keshavan, A., & Rokem, A. (2018). A browser-based tool for visualization and analysis of diffusion MRI data. *Nature Communications*, 9(1), 940. <https://doi.org/10.1038/s41467-018-03297-7>
- Zhang, J., Wei, X., Xie, S., Zhou, Z., Shang, D., Ji, R., Yu, Y., He, F., Du, Y., Ye, X., & Luo, B. (2018). Multifunctional Roles of the Ventral Stream in Language Models: Advanced Segmental Quantification in Post-Stroke Aphasic Patients. *Frontiers in Neurology*, 9, 89. <https://doi.org/10.3389/fneur.2018.00089>
- Zhang, J., Zhong, S., Zhou, L., Yu, Y., Tan, X., Wu, M., Sun, P., Zhang, W., Li, J., Cheng, R., Wu, Y., Yu, Y., Ye, X., & Luo, B. (2021). Correlations between Dual-Pathway White Matter Alterations and Language Impairment in Patients with Aphasia: A Systematic Review and Meta-analysis. *Neuropsychology Review*. <https://doi.org/10.1007/s11065-021-09482-8>
- Zhou, S., Jin, L., He, J., Zeng, Q., Wu, Y., Cao, Z., & Feng, Y. (2018). Distributed performance of white matter properties in chess players: A DWI study using automated fiber quantification. *Brain Research*, 1700, 9–18. <https://doi.org/10.1016/j.brainres.2018.07.003>