

1 DISSOCIATION OF WHITE MATTER BUNDLES IN DIFFERENT RECOVERY MEASURES IN POST-
2 STROKE APHASIA

3 Alberto Osa García, (MD, MSc)^{1,2}, Simona Maria Brambati, (PhD)^{2, 3,4}, Amélie Brisebois (MSc)^{1,2},
4 Bérengère Houzé (PhD)^{3,4} Christophe Bedetti (MSc)^{3,4}, Alex Desautels (MD, PhD)^{2,5,6}, Karine Marcotte
5 (PhD)^{1,2}

6 1 École d'orthophonie et d'audiologie, Université de Montréal, Montreal;

7 2 Centre de recherche du Centre intégré universitaire de santé et de services sociaux du Nord-de-l'île-
8 de-Montréal;

9 3 Centre de recherche de l'Institut Universitaire de Gériatrie de Montréal, Montreal;

10 4 Département de Psychologie, Université de Montréal, Montreal ;

11 5 Département de neurosciences, Université de Montréal, Montreal,

12 6 Centre d'Études Avancées en Médecine du Sommeil, Montreal

13 Corresponding author:

14 Karine Marcotte, PhD (karine.marcotte@umontreal.ca),

15 5400, Boul. Gouin Ouest, Office e-0330

16 Montréal, H4J 1C5

25

26

27

ABSTRACT

28

29

30

31

32

33

Background: Post-stroke aphasia (PSA) recovery shows high variability across individuals and at different moments during recovery. Although diffusion biomarkers from the ventral and dorsal streams have demonstrated strong predictive power for language outcomes, it is still unclear how these biomarkers relate to the various stages of PSA recovery. In this study, we aim to compare diffusion metrics and language measures as predictors of language recovery in a longitudinal cohort of participants with PSA.

34

35

36

37

38

39

40

Methods: Twenty-four participants (mean age = 73 years, 8 women) presenting PSA were recruited in an acute stroke unit. Participants underwent diffusion MRI scanning and language assessment within 3 days (acute phase) after stroke, with a behavioral follow-up at subacute (10 ± 3 days) and chronic phases (> 6 months). We used regression analyses on language performance (cross-sectional) and Δ scores at subacute and chronic timepoints (difference between acute and subacute, and subacute and chronic respectively), with language baseline scores, diffusion metrics from language-related white matter tracts, lesion size and demographic predictors.

41

42

43

44

45

46

Results: Best prediction model of performance scores used axial diffusivity (AD) from the left arcuate fasciculus (AF) in both subacute ($R^2 = 0.785$) and chronic timepoints ($R^2 = 0.626$). Moreover, prediction of change scores depended on AD from left inferior frontal-occipital fasciculus (IFOF), in subacute stage ($R^2 = 0.5$), and depended additionally on AD from right IFOF in the chronic stages ($R^2 = 0.68$). Mediation analyses showed that lesion load of left AF mediated the relationship between AD from left AF and chronic language performance.

47

48

49

50

51

Conclusion: Language performance in subacute and chronic timepoints depends on the integrity of left AF, whereas Δ scores of subacute and chronic phases depends on left IFOF, showing a dissociation of the white matter pathways regarding language outcomes. These results support the hypothesis of a functional differentiation of the dual-stream components in PSA recovery.

52

53 INTRODUCTION

54 Aphasia is one of the most devastating post-stroke cognitive sequelae, with approximately 30% of
55 stroke survivors presenting persistent language impairments (1). Despite its common prevalence,
56 predicting outcomes for individuals with acute post-stroke aphasia (PSA) remains challenging due to
57 heterogeneity of recovery patterns at the individual level (2). Recent reviews agree that lesion-related
58 factors, such as initial aphasia severity, lesion size, and affected structures (3–5) are more relevant
59 predictors of outcomes than demographic factors (e.g., age, education). Along with improvement of
60 neuroimaging techniques and statistical approaches (6–9), modeling of PSA recovery has become an
61 important goal in aphasia research. Two aspects of recovery have increasingly received attention: how
62 to measure the recovery phenomenon and what brain regions and structures are associated with this
63 recovery.

64 The time-bracketing of recovery is crucial when investigating longitudinal aphasia recovery, but its
65 definition varies among researchers. The lack of precise definition for early and late recovery periods in
66 literature may impact the interpretation of neuroimaging measurements, particularly in capturing
67 temporal pathophysiological changes in early stages of recovery (10,11), or ongoing brain vascular
68 damage in later stages (12). The groundbreaking neuroimaging findings by Saur and colleagues (12)
69 suggest three important timepoints to define recovery, namely the acute (14–16), subacute (17,18), and
70 chronic (18,19) phases. When analyzing the predictors of aphasia recovery, the choice of the outcome
71 measures is also a complex question because most severity measures are weighted by initial severity.
72 Cross-sectional language scores reflect performance at specific time points, while change scores
73 capture performance dynamics between two timepoints, representing, namely, early (acute-subacute) or
74 late (subacute-chronic) recovery phases (10–12).

75 Recent theories on aphasia recovery underscore the importance of quiescent areas in the brain's
76 post-stroke neuroplasticity (21). Significant recovery occurs both in cases of severe language
77 impairment (22,23) and in lesions affecting core cortical language areas (3,13). This suggests there
78 could be a post-stroke reorganization of the language network, which may be mediated by white matter
79 bundles. Namely, the integrity of bundles from the dorsal stream, such as arcuate fasciculus (AF) and

80 superior longitudinal fasciculus (SLF), has been associated with a better outcome in both early and
81 chronic phases (15,16,24,25). Other studies have highlighted the significant role of ventral stream
82 bundles, such as inferior fronto-occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF) and
83 uncinate fasciculus (UF), in different phases of aphasia recovery (18,26). For example, Zavanone and
84 colleagues observed a transition of the relation between language outcomes and lesioned voxels at
85 white matter bundles, shifting from more anterior-frontal in early phases of recovery to posterior-
86 temporal in later stages (18). This shift was interpreted as a progression from solely anterior parts of the
87 dorsal stream to a convergence of dorsal and ventral streams in later recovery. It was hypothesized that
88 the involvement of both dorsal and ventral streams is crucial for an optimal recovery. However, it
89 remains unclear whether overall recovery is related to preserved bundles from both streams, and
90 whether this relation is consistent across different recovery phases (21,25).

91 This study has two objectives. Firstly, we aim to compare language outcomes between early and
92 chronic phases in PSA and quantify the magnitude of change between these measures. Building on our
93 previous research (24), we anticipate an improvement in language abilities over time, albeit this
94 magnitude being similar between early and chronic phases. Secondly, we seek to explore the distinct
95 roles of dorsal and ventral stream bundle integrity in predicting language performance and changes at
96 various post-stroke timepoints. We hypothesize that if there is a specific phase-pathway association, we
97 should observe a dissociation between pathway measures and language outcomes at different
98 timepoints.

100 METHODOLOGY

102 1. Participant recruitment and procedure

103 We recruited 39 participants with aphasia resulting from a first ischemic stroke in the left middle cerebral
104 artery territory at an acute care hospital affiliated with the Centre de recherche du Centre intégré
105 universitaire de santé et de services sociaux du Nord-de-l'Île-de-Montréal (CIUSSS NIM). Eligibility,
106 confirmed by a neurologist, required proficient French and/or English skills before the stroke, with
107 exclusion criteria including left-handedness, major psychiatric or developmental disorders, severe

108 perceptual deficits, or other major neurological conditions. Following the protocol outlined by Saur and
109 colleagues (13), language assessments and MRI scans were conducted within 3 days (acute), 7-15
110 days (early subacute), and at least 6 months (chronic) post-stroke. Fifteen participants completed only
111 one or two assessments and/or withdrew from the study. Among the remaining participants, only 21
112 agreed to a subacute MRI scan, and of these, only 15 consented to a chronic MRI scan. To ensure
113 consistency across timepoints, we utilized data from acute MRI scans alongside language data from all
114 three timepoints, resulting in a final sample of 24 participants (mean age= 70 years, SD= 13; 8 women).
115 Ethics approval was granted by the CIUSSS NIM ethics committee (#MP-32-2018-1478), and written
116 informed consent was obtained from all participants or their legally authorized representatives.

117 118 2. Language assessment and variable creation

119 Based on previous studies (22,24), we calculated composite scores (CS), representing cross-
120 sectional language performance, consisting of three sub-scores: comprehension, repetition and naming.
121 The comprehension sub-score was composed of the word-sentence comprehension score from the
122 Montreal-Toulouse aphasia battery (27) and the score of the revised short version of the Token Test (28).
123 The repetition sub-score was composed of the word repetition and sentence repetition tasks from the
124 MT-86 (27). The naming sub-score was composed of the semantic fluency score of the Protocole
125 Montréal d'Évaluation de la Communication (29), and either the score of the Dénomination orale
126 d'images (30) for participants speaking French or the score of the Boston Naming Test (31) for
127 participants speaking English. All these tests are extensively used in both English and French speech
128 pathology practice. The total score of each sub-score was reported on 10, resulting in CS maximum = 30.
129 Three CS's were calculated for each participant: acute (CS₁), subacute (CS₂), and chronic (CS₃) scores.

130 Then we calculated the change of these scores in time. Change measures included early change
131 ($\Delta CS_{1-2} = CS_2 - CS_1$) and late change ($\Delta CS_{2-3} = CS_3 - CS_2$) scores. Additionally, we computed relative
132 recovery scores, with early relative change score ($r\Delta CS_{1-2} = (CS_2 - CS_1) / CS_1$) and late relative change
133 score ($r\Delta CS_{2-3} = (CS_3 - CS_2) / CS_2$), emphasizing the gains of individuals with more severe impairments
134 at baseline (32).

135 3. Neuroimaging procedure

3.1. Image acquisition

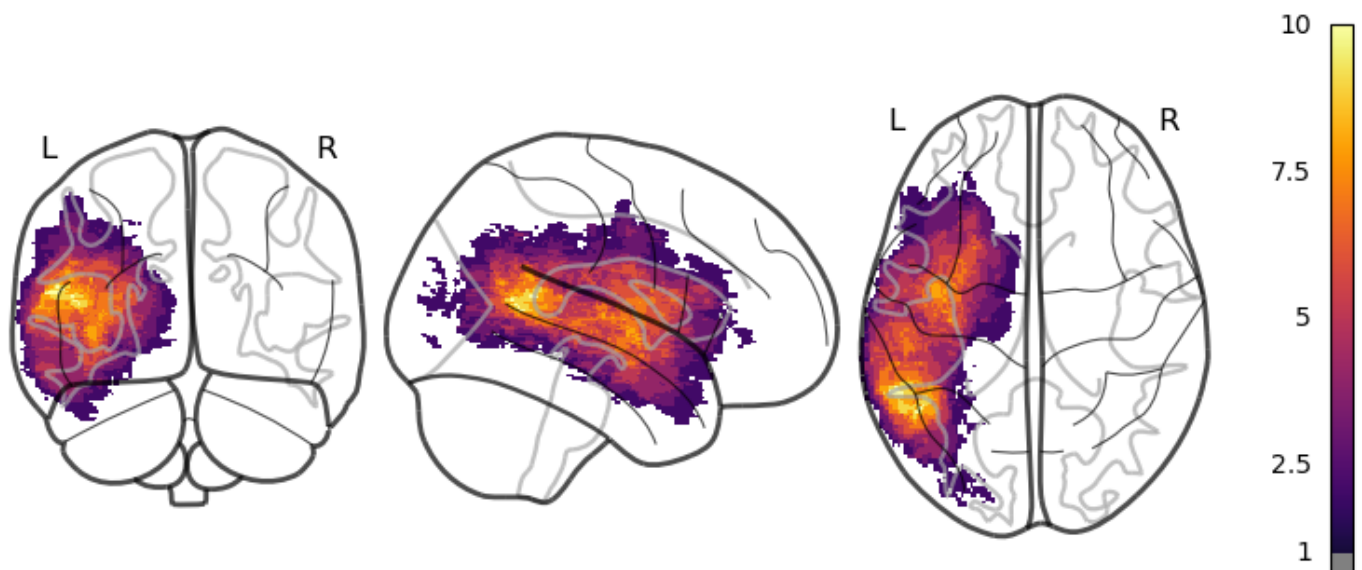
An MRI scan including a diffusion sequence was performed on each participant on the day of the initial language assessment. The MRI images were acquired with a Skyra 3T scanner (Siemens Healthcare, USA). A high-resolution 3D T1-weighted image was acquired (TR = 2200 ms, TE = 2.96 ms, TI = 900 ms, FOV = 250 mm, voxel size = 1x1x1 mm³, matrix = 256x256, 192 slices, flip-angle = 8°), in a Magnetization Prepared Rapid Gradient Echo (MP-RAGE) sequence. The MRI diffusion-weighted images had the following parameters: 65 images with non-collinear diffusion direction in b = 1000 s/mm², posterior-anterior acquisition (TR = 11000 ms, TE = 86 ms, field of view = 230 mm, voxel resolution = 2 x 2 x 2 mm³, flip angle = 90° , bandwidth = 1700 , EPI factor = 67), and two T2-weighted images with b = 0 s/mm², one being a posterior-anterior acquisition, the other an anterior-posterior acquisition (time of acquisition = 12 min 30 s).

3.2. DWI pre-processing and metric extraction

All pre-processing corrections and diffusion measure extractions were completed with the automated and reproducible pipeline Tractoflow (33). The procedure was supervised by three co-authors (CB, BH and SMB). First, we performed noise correction, using Marchenko-Pastur principal component analysis through MRtrix3 (34); subject movement and induced distortion correction through the FSL package (35); and a correction for N4 bias through ANTs package (36). Then, we performed the DTI metric extraction using DIPY (37), and T1-weighted image registration using non-linear SyN through ANTs (36). We used the Atlas Based Segmentation profile of TractoFlow version 2.3.0, using Freesurfer anatomical images in the within-subject template space, particularly recommended on pathological data (33). Once the tractogram was reconstructed, diffusion measures, namely fractional anisotropy (FA), mean diffusivity (MD) and axial diffusivity (AD), were extracted from bundles of ventral stream (IFOF, ILF and UF) and the dorsal stream (AF and SLF) bilaterally.

Finally, we performed a semi-automated segmentation of each brain lesion in acute phase using *Clusterize* (38). This toolbox (<http://www.medizin.uni-tuebingen.de/kinder/en/research/neuroimaging/software/>), running with SPM12 under Matlab (The Mathworks, Inc., Natick, MA), has previously shown to have a good reliability in lesion demarcation in acute post-stroke patients (De Haan et al., 2015). Clusters of hypointense voxels were first identified on

164 MD maps (set with default parameters). Then, cluster(s)-of-interest were manually selected and adjusted
165 to fit the lesion in each slice and adjusted (if needed) with MI-brain software (Imeka Solutions Inc.), with
166 the help of MD maps and b0 DWI maps by a study team member (BH). Finally, each lesion file was
167 counter-verified by another team member (SMB) with the help of MD maps and b0 DWI maps. Both
168 raters are experiences in lesion delineation and were blinded to the participant's identity. Volume of the
169 lesions and intracranial volumes were extracted in mL and corrected for intracranial volume. Lesion
170 loads for each left-hemisphere tract were also calculated using the mask of the lesion and the parceled
171 area of the tracts, with a correction for tract volume afterwards.



172

173

174 Figure 1. Lesion overlay map of participants in the study. Color scale indicates minimum number of
175 participants with the same location for lesion voxels.

176

177

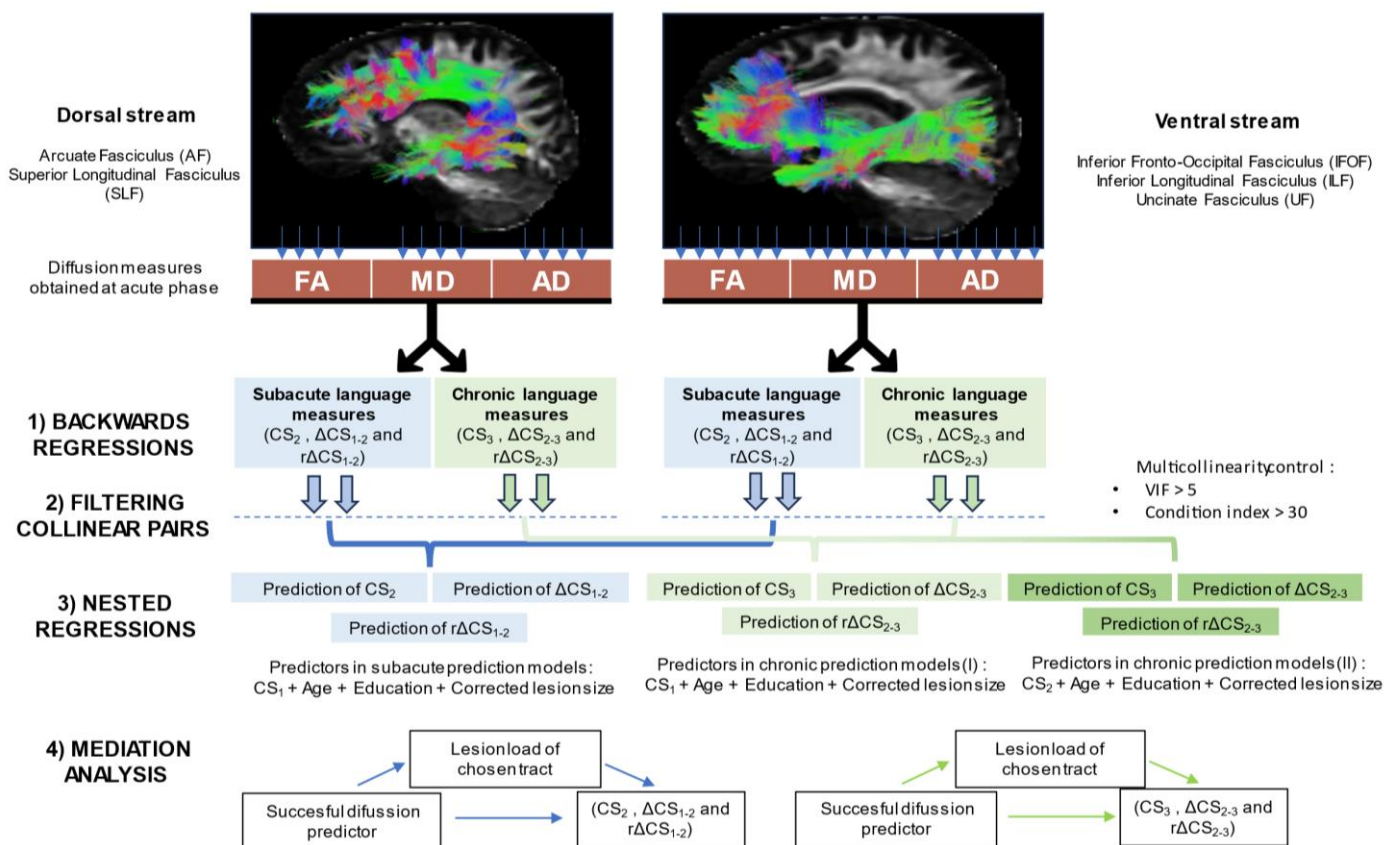
178 3.3. Statistical proceeding

179

180 To evaluate the impact of time on language performance scores, we conducted a one-way ANOVA
181 on all CS, and paired tests on the difference of Δ CS and $r\Delta$ CS in both early and late recovery phases.
182 Additionally, we examined whether receiving rt-PA affected any language outcome. Next, we identified
183 diffusion predictors for inclusion in regression models. In the subacute phase, the dependent variables

184 were CS_2 , ΔCS_{1-2} and $r\Delta CS_{1-2}$. In the chronic phase, the dependent variables were CS_3 , ΔCS_{2-3} and
185 $r\Delta CS_{2-3}$. All analyses were performed using SPSS (version 28.0.1.0). Model selection is illustrated in
186 Figure 1 and involved the following steps:

- 187 1) For each stream (dorsal and ventral stream), we conducted backwards regressions, entering the
188 diffusion measures of each representative tract (dorsal stream: AF, SLF; ventral stream: ILF, IFOF, UF)
189 from both hemispheres. Collinearity was monitored, and predictors were retained based on Variance
190 Inflation Factor (VIF) and Condition Indices (CI). Predictors were excluded if they exhibited a correlation >
191 0.6 with another predictor in the same final model, showing a $VIF > 3$ or $CI > 30$. In the case of
192 collinearity, the predictor with the higher standardized beta coefficient was retained.
- 193 2) We conducted nested linear regressions using retained predictors from step 1, incorporating lesion,
194 language, and demographic measures. Each regression comprised a null model with only diffusion
195 measures predicting each dependent variable, and a complete model with diffusion measures and
196 covariates (age, education, lesion size, and the CS of the timepoint immediately prior to the studied
197 timepoint). Subacute models included CS_1 , while chronic models had two versions, one with CS_1 and
198 another with CS_2 as the baseline score. This comparison aimed to evaluate the relationship between
199 acute (initial), subacute, and chronic scores, considering an expected larger change in the early weeks
200 of recovery.
- 201 3) A-posteriori mediation analyses were conducted to assess the influence of lesion loads in the left
202 hemisphere language bundles for each successful measure in the nested models.
203 Each mediation analysis involved a diffusion measure as the predictor and language measures as
204 responses.



205

206

Figure 2. Procedure used for variable selection and model construction.

207

CS= Composite score; FA= Fractional Anisotropy; MD= Mean Diffusivity; AD: Axial Diffusivity.

208

209

RESULTS

210

Table 1. Demographic and clinical characteristics of participants

211

Variable	Mean \pm SD	Range
Sex (M/F)	16 M; 8 F	
Age (years)	70 \pm 13	(47, 95)
Education (years)	11.3 \pm 5.5	(6, 26)
Lesion site (F/T/P/SC)*	11/10/10/12	
Lesion size (mL)	24.75 \pm 24.7	(0.16, 84.42)
Thrombolysis received (Y/N)	13/11	
Time Stroke - Assessment ₁ (days)	2 \pm 0.8	(1, 3)
Time Stroke - Assessment ₂ (days)	10 \pm 2.1	(7, 15)
Time Stroke - Assessment ₃ (days)	270 \pm 60	(180, 500)
Speech / Language therapy (Y/N) †	21/3	
CS ₁	15.3 \pm 9.13	(0, 28.5)
CS ₂	20.3 \pm 8.5	(0, 29.5)
CS ₃	25.8 \pm 5.3	(5.3, 29.7)
Δ CS ₁₋₂	3.8 \pm 5.2	(-9.9, 16.5)
Δ CS ₂₋₃	5.07 \pm 5.8	(-0.56, 18.12)
r Δ CS ₁₋₂ ‡	22.37 / 3.8 + 101.1 / 3.11	(-46.1, 372.7) / (-4.52, 6.61)
r Δ CS ₂₋₃ ‡	13.55 / 3.26 + 59.7 / 2.46	(-2.84, 219.36) / (-1.76, 6.08)

*F= Frontal; T= Temporal; P= Parietal; SC= Subcortical

† Y = Received speech/language therapy; N = Did not receive therapy. Minimum reported = one month with three weekly sessions with a speech-language therapist, maximum reported duration = 12 months.

‡ Raw relative change scores / Relative change scores after Box-Cox transformation to address high heteroscedasticity. Analyses were performed using transformed values.

Demographic, clinical characteristics and CS are reported in Table 1. Friedman's ANOVA analysis across the three CS was significant ($\chi^2 = 28.2$, $p < 0.001$, effect size Kendall's $W = 0.587$). Conover post-hoc comparisons, with Holm correction applied, showed a difference between the acute and subacute timepoints (T [46] = 2.32, $p = 0.025$) and between the subacute and chronic timepoints (T [46] = 5.297, $p < 0.001$). No difference in change measures were found between Δ CS₁₋₂ and Δ CS₂₋₃ (t [23] = -0.724, $p = 0.477$) or between r Δ CS₁₋₂ and r Δ CS₂₋₃ ($W = 169$, $p = 0.6$). Participants that received rt-PA had a higher CS₁ (t [22] = -2.225, $p = 0.037$, Cohen's $d = 0.45$), CS₂ ($W = 37$, $p = 0.047$, Cohen's $d = 0.23$) and CS₃ ($W = 36$, $p = 0.041$, Cohen's $d = 0.23$) than their counterparts, but there was no difference for Δ CS₁₋₂ ($W = 72$, $p = 0.9$), and Δ CS₂₋₃ ($W = 96$, $p = 0.16$). Visualization of the trajectory of sub-scores, CS and change scores can be found in the supplementary material.

DIFFUSION PREDICTORS

231

232

233

234

235

236

237

238

239

240

COMPLETE REGRESSION MODELS

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

A POSTERIORI-MEDIATION ANALYSES

257 We only conducted mediation analyses between the lesion load of L-AF_{AD}, L-IFOF_{AD} and the
258 response measures that were predicted in each model, since only those two measures proved to be
259 significant in the full models. No mediation effect was found in the prediction of raw or relative change
260 measures using L-IFOF_{AD}. No mediation effect was either found in the prediction of CS₂ using L-AF_{AD} ,
261 however we found that the lesion load of AF worked as a mediator in the prediction of CS₃ using L-AF_{AD}
262 (Indirect effect = 21566, average mediation effect = 60%, CI= [4738.08, 44294.83], proportional CI=
263 [0.17,1.4]; p = 0.008). In presence of the lesion load of AF, L-AF_{AD} was not a significant predictor of CS₃
264 (β = 0.202, p=0.255), whereas the lesion load of AF proved to be significant (β = - 0.616, p=0.002).
265 Visualization of the mediation effect of Lesion load on L-AF_{AD} regression on CS₃ can be found in the
266 supplementary material.

267

Table 2. Results from backwards regression models for each dependent variable. Only predictors with significant coefficients at $\alpha=0.05$ and individual R^2

268

obtained from part correlations are reported. Highly collinear predictor pairs were excluded.

	Subacute recovery phase									Chronic recovery phase					
	CS ₂			$\Delta CS_{1,2}$			r $\Delta CS_{1,2}$			CS ₃			$\Delta CS_{2,3}$		
	Predictor	R ²	β_s	Predictor	R ²	β_s	Predictor	R ²	β_s	Predictor	R ²	β_s	Predictor	R ²	β_s
Dorsal stream	L-AF _{AD}	0.275	0.524							L-AF _{AD}	0.266	0.516			
	L-AF _{FA}	0.198	0.445							L-AF _{FA}	0.192	0.438			
Ventral stream	L-UF _{MD}	0.169	0.411	L-IFOF _{AD}	0.27	0.522	L-IFOF _{AD}	0.168	0.41				L-UF _{MD}	0.186	-0.431
													L-IFOF _{AD}	0.33	-0.81
													R-IFOF _{AD}	0.193	0.5

269

270

271

272

273

274

275 Table 3. Coefficients in prediction models of performance scores, raw change scores and relative change
 276 scores. Variables and elements in bold represent significant additions in the complete regression models at
 277 $\alpha=0.05$.

Model	CS ₂ model		CS ₃ model with CS ₁		CS ₃ model with CS ₂	
Nested model	R ² = 0.31		R ² = 0.26		R ² = 0.26	
Complete model	R ² = 0.78	R² change p < 0.001	R ² = 0.5	R ² change p = 0.06	R ² = 0.62	R² change p = 0.01
Predictors	Stand.β	P value	Stand.β	P value	Stand.β	P value
Intercept		0.85		0.425		0.36
CS ₁ /CS ₂	0.53	0.015*	0.465	0.019*	0.81	0.004*
Age	- 0.11	0.44	- 0.24	0.21	- 0.19	0.29
Education	- 0.20	0.15	0.1	0.61	0.275	0.15
Lesion size	- 0.18	0.39	0.12	0.70	0.17	0.46
L-AF _{AD}	0.11	0.54	0.50	0.034*	0.41	0.077
L-UF _{MD}	0.14	0.33	-	-	-	-
Model	ΔCS ₁₋₂ model		ΔCS ₂₋₃ model with CS ₁		ΔCS ₂₋₃ model using CS ₂	
Nested model	R ² = 0.27		R ² = 0.39		R ² = 0.39	
Complete model	R ² = 0.50	R ² change p = 0.123	R ² = 0.69	R² Change p = 0.023*	R ² = 0.68	R² Change p = 0.029*
Predictors	Stand.β	P value	Stand.β	P value	Stand.β	P value
Intercept		0.150		0.425		0.150
CS ₁ /CS ₂	- 0.737	0.017*	-0.495	0.077	- 0.737	0.017*
Age	- 0.059	0.733	0.048	0.765	- 0.059	0.733
Education	- 0.224	0.290	0.08	0.660	- 0.224	0.290
Lesion size	- 0.341	0.222	0.102	0.663	- 0.341	0.222
L-IFOF _{AD}	0.407	0.048*	- 0.65	0.026*	0.407	0.048*

R-IFOF _{AD}	-	-	0.374	0.095	-	-
L-UF _{MD}	-	-	-0.01	0.965	-	-
Model	r Δ CS ₁₋₂ model		r Δ CS ₂₋₃ model with CS ₁		r Δ CS ₂₋₃ model with CS ₂	
Nested model	R ² = 0.17					
Complete model	R ² = 0.37	R ² Change p = 0.26	R ² = 0.69		R ² = 0.62	
Predictors	Stand.β	P value	Stand.β	P value	Stand.β	P value
Intercept	-	0.306	-	0.527	-	0.187
CS ₁ /CS ₂	- 0.705	0.039*	-0.582	0.003	-0.706	0.008
Age	0.092	0.638	0.157	0.318	0.142	0.336
Education	- 0.221	0.351	0.236	0.161	0.075	0.663
Lesion size	- 0.349	0.264	0.311	0.728	0.065	0.763
L-IFOF _{AD}	0.276	0.219	-	-	-	-

278

279

280

281

282

283

284

285

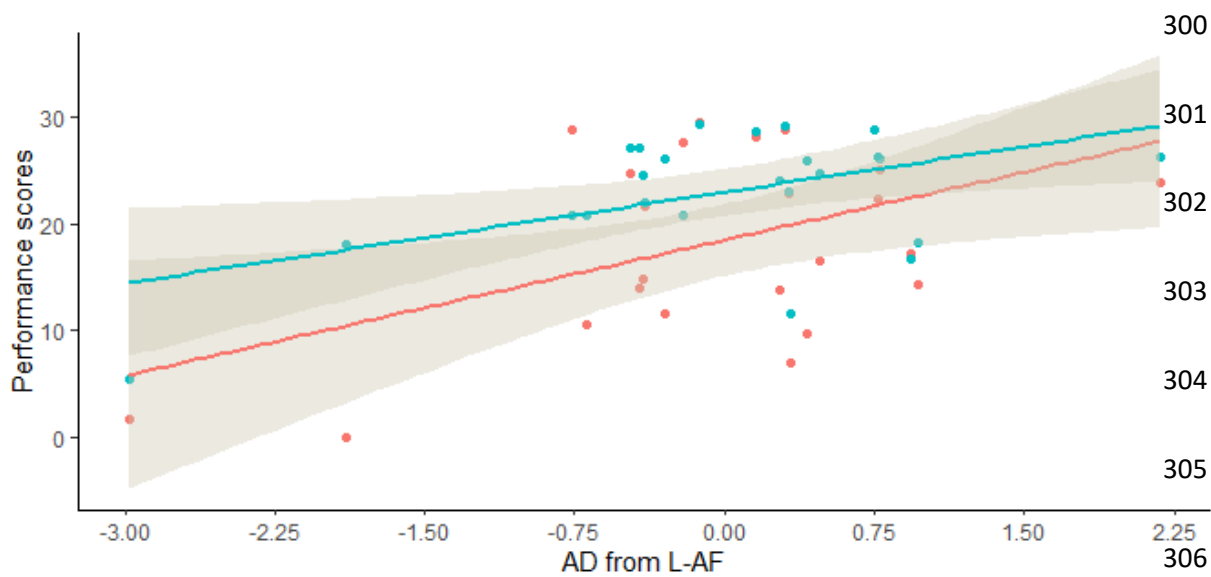
286

287

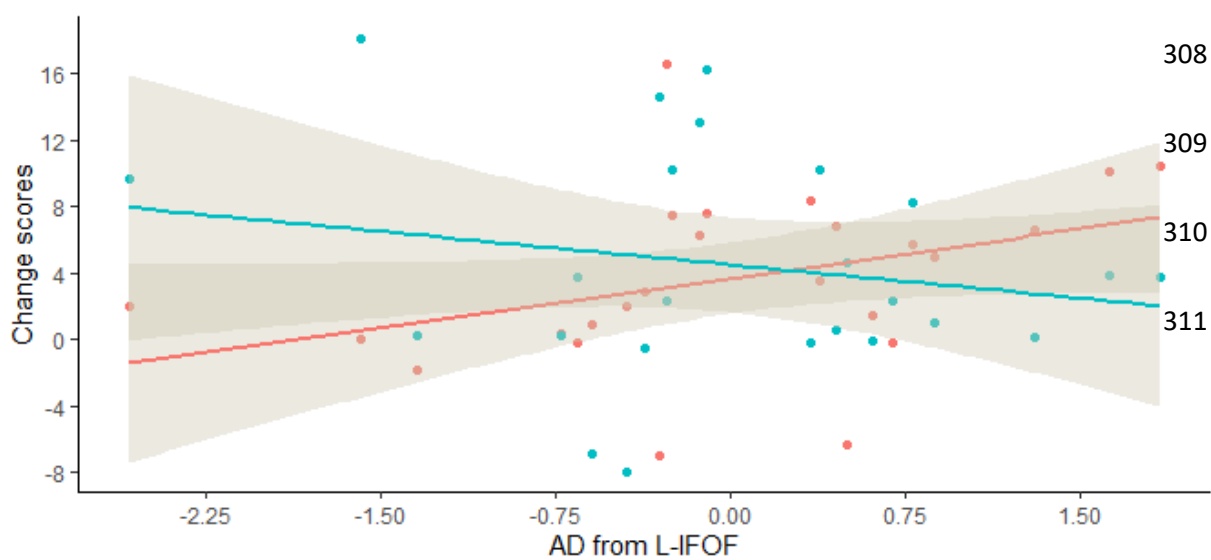
288

289

290
291
292
293
294
295
296
297
298
299



Recovery phase • Subacute • Chronic



312

313

314

315

316

317

318

319 Figure 3. Scatterplots with the relation of predictive diffusion (AD from L-AF and AD from L-IFOF)
320 measures and the language outcomes for which they had significant coefficients. X-axis variables are
321 shown scaled.

322

323 DISCUSSION

324 In this study, we first compared the language outcomes from early and late phases of recovery and
325 confirmed continuous improvement in language abilities. Consistent with seminal work in this field (23),
326 we observed a similar degree of recovery in the early and late phases. Then, we explored the
327 relationship between diffusion measures from the dual-stream system tracts and language outcomes in
328 PSA. Results showed that acute AD from left AF predicted performance scores in subacute and chronic
329 timepoints, while acute AD from left IFOF predicted early and late change scores, showing a negative
330 trend in the chronic phase. Additionally, AD from right IFOF correlated positively with chronic change
331 scores. Baseline language scores enhanced complete prediction models in all cases. All language
332 outcomes (performance and change scores) could be well predicted by corresponding baseline
333 language abilities, diffusion measures and covariates, as expected because of the relation between the
334 outcome measure and the aphasia severity (i.e., CS) from the previous phases. Finally, results showed
335 that lesion load only influenced the relationship between left AF and chronic cross-sectional outcomes.
336 In summary, it was the nature of language measures (performance vs change scores) that differentiated
337 the relationship between white matter tracts and language measures, rather than chronology of the

338 language measures. The performance measure represents time-locked performance abilities whereas
339 the change score represents the dynamic of performance between two timepoints. Thus, the present
340 results could suggest that the associated tracts (i.e. left AF, left IFOF, left UF and right IFOF) may have
341 different roles in the recovery trajectory. To our knowledge, this is the first time that such dissociation is
342 reported.

343 L-AF_{AD} effectively predicted subacute and chronic performance scores in our PSA cohort, consistent
344 with previous findings linking AF to optimal language abilities in PSA individuals, including naming (39),
345 repetition (40,41), comprehension (39) and spontaneous speech (42,43). Other studies that have tried to
346 relate change scores and AF measures have not obtained significant results either (44–46). Bae and
347 colleagues (45) demonstrated a link between left AF diffusion changes (increased FA over time) and
348 language score variations after 6 months, suggesting a positive impact of improved diffusion stability on
349 language, albeit with different measurement timing than our study. However, change of diffusion
350 measures requires a chronic measure, and we intended to investigate the utility of acute measures as
351 predictors of chronic outcomes. However, this prediction was conditional on the overall integrity of the
352 left AF, since lesion load seemed to be a mediator in that relation, echoing the importance of the left
353 hemisphere dorsal pathway over its right counterpart in recovery (39,47). The lack of a similar effect on
354 left IFOF measures and the exclusive association of left, not right, AF measures with language
355 improvement emphasizes the sensitivity of this bundle to anatomical disruptions, consistent with other
356 cohort studies (48).

357 Regarding ventral stream tracts, an exclusive relation to change scores was found in both recovery
358 phases. Two results notably stand out: AD from right IFOF significantly predicted late changes, while AD
359 from left IFOF changed the direction of its relation from positive (in early change prediction) to negative
360 (in late change prediction). These two results may reflect the bilateral nature of the ventral pathway,
361 particularly evident in the chronic phase, which would point to a requirement for bilateral ventral
362 structures in long-term recovery. It should be noted, though, that this finding does not entirely support
363 the need for right dorsal-stream structures, but rather reminds the inherent bilateral nature of the
364 language abilities supported by the ventral streams. A recent study, exploring right hemisphere
365 connectivity, precisely suggested a crucial role of these structures in later stages compared to early
366 stages (17). Our results underscore this claim, highlighting the significance of language-related tracts

367 from the dual-stream model. We consider important to remind how early phases (i.e., subacute
368 performance, or early changes between subacute and acute) are indeed predictable, and our data
369 suggest that higher stability of white matter bundles from both streams support language from the very
370 beginning of the recovery process, being proof that early recovery can be well captured with
371 neurobiological measures.

372 Our findings also primarily indicate a relationship between AD measures and language recovery,
373 with FA only emerging as a predictor in one instance. Interestingly, almost all our regression analyses
374 consistently identified AD as a better predictor. While AD has received less attention than FA in aphasia
375 literature (15,39,46), recent work focusing on the late subacute phase in stroke survivors demonstrated
376 its association with positive language outcomes (49). Higher AD in the language-related tracts could
377 reflect higher stability, which also tends to persist over time. However, it is usual that AD rises in the
378 acute phase of recovery, whereas FA remains stable for longer periods. As suggested by Moulton and
379 colleagues (49), AD may provide valuable information on in the acute prognostic, since lower AD could
380 indicate poor hyperacute recovery after stroke, though it may not decisively predict recovery in later
381 stages.

382 Our study has certainly its limitations. Beyond the modest size of the sample, we relied on structural
383 data to explore neural correlates of recovery, whereas connectivity measures can provide a deeper
384 insight on the status of each tract and in each phase. Nevertheless, we consider that this study is
385 valuable for the importance of reporting more than two timepoints in the same cohort including early
386 acute measures, which has yielded important evidence about recovery in recent literature (2,13,44).
387 Another limitation is the absence of specific cortical measures, although literature has vastly shown
388 utility of both white and grey matter measures in predicting recovery (15,41). The choice of variables is
389 also an important issue in aphasia recovery studies. While score changes are increasingly utilized in
390 aphasia studies to represent language gains (2,17,21,44) they may not fully consider individual baseline
391 abilities, potentially affecting the interpretation of results (32). Bae and colleagues (50) also addressed
392 this concern but did not find significant results using proportional recovery measures. There may not be
393 a linear relationship between relative recovery measures and diffusion measures (possibly due to
394 heteroscedasticity), but we consider that behavioral studies should rightfully assess individual recovery
395 patterns to enable meaningful group-level comparisons.

396 In conclusion, our data suggest that bundles from the dual-stream model may have a role in the
397 recovery process, possibly either mediated by the language processes they support, or by a
398 compensation that they confer for the remaining language abilities after the disruption of the language
399 network. These hypotheses should be tested in future venues to further elucidate the nature of PSA
400 recovery.

401 402 Acknowledgments

403 We thank all participants, their families and health professionals from the clinical setting for their
404 collaboration all along the project.

405 406 Sources of Funding

407 This project was funded by a grant-in-aid from the Heart and Stroke Foundation (G-16-00014039
408 and G-19-0026212) as well as a project grant from the Canadian Institutes of Health Research (CIHR
409 470371) to K.M., A.D. and S.M.B. K.M. holds a Career Award from the "Fonds de Recherche du
410 Québec – Santé". A.O.G. holds a "Fonds de Recherche du Québec – Santé" scholarship.

411 412 Disclosures

413 AD received research grants from Canopy Growth and Eisai, served on scientific advisory boards for
414 Eisai, Jazz Pharma and Eisai, as well as honoraria from speaking engagements from Jazz Pharma and
415 Paladin Labs. The remaining authors do not have any conflict of interest to declare.

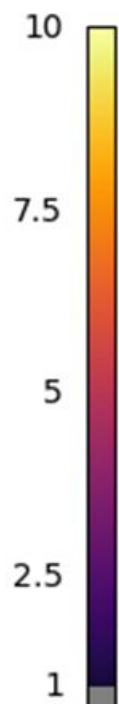
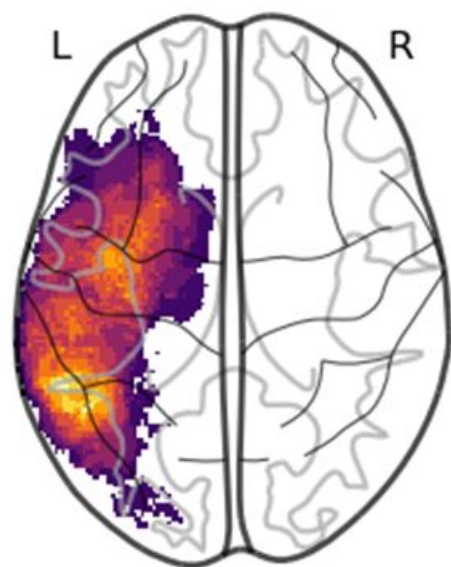
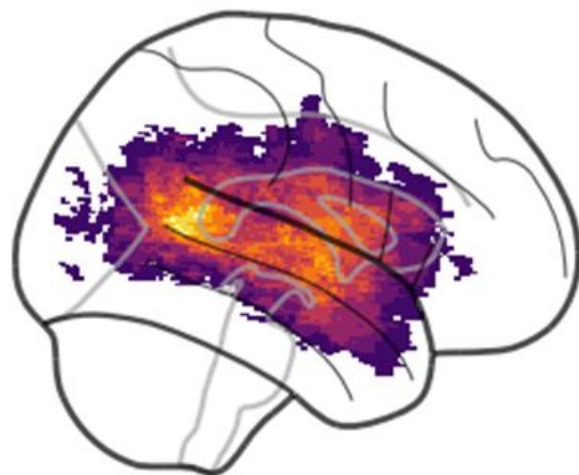
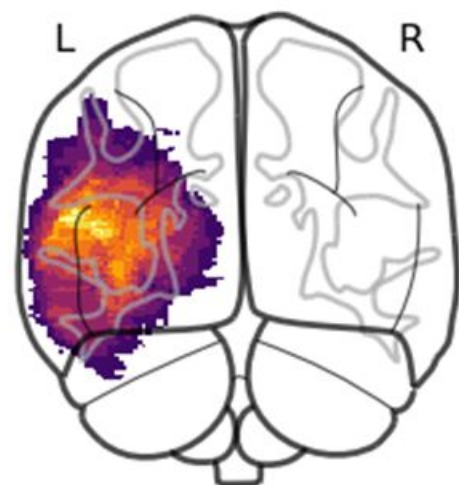
416 417 REFERENCES

- 418 1. Frederick A, Jacobs M, Adams-Mitchell CJ, Ellis C. The Global Rate of Post-Stroke Aphasia. *Perspect ASHA Spec*
419 *Interest Groups*. 2022 Oct 20;7(5):1567–72.
- 420 2. Wilson SM, Entrup JL, Schneck SM, Onuscheck CF, Levy DF, Rahman M, et al. Recovery from aphasia in the first
421 year after stroke. *Brain*. 2023 Mar 1;146(3):1021–39.

- 422 3. Flowers HL, Skoretz SA, Silver FL, Rochon E, Fang J, Flamand-Roze C, et al. Poststroke Aphasia Frequency,
423 Recovery, and Outcomes: A Systematic Review and Meta-Analysis. In: Archives of Physical Medicine and
424 Rehabilitation. 2016.
- 425 4. Plowman E, Hentz B, Ellis C. Post-stroke aphasia prognosis: A review of patient-related and stroke-related factors.
426 J Eval Clin Pract. 2012;18(3):689–94.
- 427 5. Watila MM, Balarabe B. Factors predicting post-stroke aphasia recovery. J Neurol Sci. 2015;
- 428 6. Kristinsson S, Zhang W, Rorden C, Newman-Norlund R, Basilakos A, Bonilha L, et al. Machine learning-based
429 multimodal prediction of language outcomes in chronic aphasia. Hum Brain Mapp. 2021 Apr 15;42(6):1682–98.
- 430 7. Schumacher R, Halai AD, Lambon Ralph MA. Assessing and mapping language, attention and executive
431 multidimensional deficits in stroke aphasia. Brain. 2019 Oct 1;142(10):3202–16.
- 432 8. Halai AD, Woollams AM, Lambon Ralph MA. Using principal component analysis to capture individual differences
433 within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural
434 correlates of speech fluency, phonology and semantics. Cortex. 2017 Jan;86:275–89.
- 435 9. Iorga M, Higgins J, Caplan D, Zinbarg R, Kiran S, Thompson CK, et al. Predicting language recovery in post-stroke
436 aphasia using behavior and functional MRI. Sci Rep. 2021 Apr 19;11(1):8419.
- 437 10. Bernhardt J, Hayward KS, Kwakkel G, Ward NS, Wolf SL, Borschmann K, et al. Agreed definitions and a shared
438 vision for new standards in stroke recovery research: The Stroke Recovery and Rehabilitation Roundtable
439 taskforce. Int J Stroke. 2017 Jul;12(5):444–50.
- 440 11. Osa García A, Brambati SM, Desautels A, Marcotte K. Timing stroke: A review on stroke pathophysiology and its
441 influence over time on diffusion measures. J Neurol Sci. 2022 Oct;441:120377.
- 442 12. Kalaria RN, Akinyemi R, Ihara M. Stroke injury, cognitive impairment and vascular dementia. Biochim Biophys
443 Acta BBA - Mol Basis Dis. 2016 May;1862(5):915–25.
- 444 13. Saur D, Lange R, Baumgaertner A, Schraknepper V, Willmes K, Rijntjes M, et al. Dynamics of language
445 reorganization after stroke. Brain. 2006 Jun;129(Pt 6):1371–84.
- 446 14. Zhu Y, He Y, Von Deneen KM, Zhang M. Abnormal Integrity of White Matter Tracts in Acute Lacunar Stroke
447 Patients. Brain Impair. 2019 Mar;20(1):49–57.
- 448 15. Yang M, Yang P, Fan YS, Li J, Yao D, Liao W, et al. Altered Structure and Intrinsic Functional Connectivity in Post-
449 stroke Aphasia. Brain Topogr. 2018;31(2):300–10.
- 450 16. Zhang J, Wei X, Xie S, Zhou Z, Shang D, Ji R, et al. Multifunctional Roles of the Ventral Stream in Language Models:
451 Advanced Segmental Quantification in Post-Stroke Aphasic Patients. Front Neurol. 2018 Feb 27;9:89.
- 452 17. Sihvonen AJ, Vadinova V, Garden KL, Meinzer M, Roxbury T, O'Brien K, et al. Right hemispheric structural
453 connectivity and poststroke language recovery. Hum Brain Mapp. 2023 May;44(7):2897–904.
- 454 18. Zavanone C, Samson Y, Arbizu C, Dupont S, Dormont D, Rosso C. Critical brain regions related to post-stroke
455 aphasia severity identified by early diffusion imaging are not the same when predicting short- and long-term
456 outcome. Brain Lang. 2018 Nov;186:1–7.
- 457 19. Stockert A, Wawrzyniak M, Klingbeil J, Wrede K, Kümmerer D, Hartwigsen G, et al. Dynamics of language
458 reorganization after left temporo-parietal and frontal stroke. Brain. 2020 Mar 1;143(3):844–61.
- 459 20. Van Oers CAMM, Van Der Worp HB, Kappelle LJ, Raemaekers MAH, Otte WM, Dijkhuizen RM. Etiology of
460 language network changes during recovery of aphasia after stroke. Sci Rep. 2018 Jan 16;8(1):856.

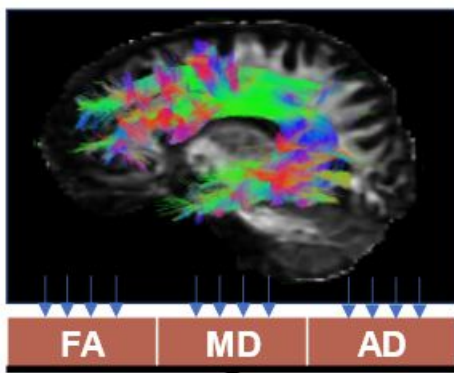
- 461 21. Stefaniak JD, Halai AD, Lambon Ralph MA. The neural and neurocomputational bases of recovery from post-
462 stroke aphasia. *Nat Rev Neurol*. 2020 Jan;16(1):43–55.
- 463 22. Lazar RM, Minzer B, Antonello D, Festa JR, Krakauer JW, Marshall RS. Improvement in aphasia scores after
464 stroke is well predicted by initial severity. *Stroke*. 2010;41(7):1485–8.
- 465 23. Pedersen PM, Stig Jørgensen H, Nakayama H, Raaschou HO, Olsen TS, Jørgensen HS, et al. Aphasia in acute
466 stroke: Incidence, determinants, and recovery. *Ann Neurol*. 1995;38(4):659–66.
- 467 24. Osa García A, Brambati SM, Brisebois A, Désilets-Barnabé M, Houzé B, Bedetti C, et al. Predicting Early Post-
468 stroke Aphasia Outcome From Initial Aphasia Severity. *Front Neurol*. 2020 Feb 21;11:120.
- 469 25. López-Barroso D, De Diego-Balaguer R. Language Learning Variability within the Dorsal and Ventral Streams as a
470 Cue for Compensatory Mechanisms in Aphasia Recovery. *Front Hum Neurosci*. 2017 Sep 27;11:476.
- 471 26. Yang M, Li Y, Li J, Yao D, Liao W, Chen H. Beyond the Arcuate Fasciculus: Damage to Ventral and Dorsal Language
472 Pathways in Aphasia. *Brain Topogr*. 2017 Mar;30(2):249–56.
- 473 27. Nespoulous JL, Lecours AR, Lafond D. MT-86-Protocole Montréal-Toulouse d'examen linguistique de l'aphasie.
474 Ortho-Ed Isbergues. 1986;
- 475 28. De Renzi E, Faglioni P. Normative data and screening power of a shortened version of the Token Test. *Cortex*.
476 1978;14(1):41–9.
- 477 29. Joannette Y, Ska B, Côté H. Protocole Montréal d'Évaluation de la Communication. Isbergues, France: Ortho
478 Édition; 2004.
- 479 30. Deloche G, Hannequin D. Test de dénomination orale d'images: DO-80. Éditions du centre de psychologie
480 appliquée; 1997.
- 481 31. Kaplan E, Goodglass H, Weintraub S. Boston Naming Test. Philadelphia: Lea & Febiger; 1983.
- 482 32. Ali M, VandenBerg K, Williams LJ, Williams LR, Abo M, Becker F, et al. Predictors of Poststroke Aphasia Recovery:
483 A Systematic Review-Informed Individual Participant Data Meta-Analysis. *Stroke*. 2021 May;52(5):1778–87.
- 484 33. Theaud G, Houde JC, Boré A, Rheault F, Morency F, Descoteaux M. TractoFlow: A robust, efficient and
485 reproducible diffusion MRI pipeline leveraging Nextflow & Singularity. *NeuroImage*. 2020 Sep;218:116889.
- 486 34. Tournier JD, Smith R, Raffelt D, Tabbara R, Dhollander T, Pietsch M, et al. MRtrix3: A fast, flexible and open
487 software framework for medical image processing and visualisation. *NeuroImage*. 2019 Nov;202:116137.
- 488 35. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *NeuroImage*. 2012;62(2):782–90.
- 489 36. Avants B, Tustison NS. Advanced Normalization Tools; ANTS. 2009.
- 490 37. Garyfallidis E, Brett M, Amirbekian B, Rokem A, van der Walt S, Descoteaux M, et al. Dipy, a library for the
491 analysis of diffusion MRI data. *Front Neuroinform*. 2014;8:8.
- 492 38. Clas P, Groeschel S, Wilke M. A Semi-Automatic Algorithm for Determining the Demyelination Load in
493 Metachromatic Leukodystrophy. *Acad Radiol*. 2012;19(1):26–34.
- 494 39. Ivanova MV, Isaev DY, Dragoy OV, Akinina YS, Petrushevskiy AG, Fedina ON, et al. Diffusion-tensor imaging of
495 major white matter tracts and their role in language processing in aphasia. *Cortex*. 2016 Dec;85:165–81.
- 496 40. Berthier ML, Lambon Ralph MA, Pujol J, Green C. Arcuate fasciculus variability and repetition: The left
497 sometimes can be right. *Cortex*. 2012 Feb;48(2):133–43.

- 498 41. Breier JJ, Hasan KM, Zhang W, Men D, Papanicolaou AC. Language dysfunction after stroke and damage to white
499 matter tracts evaluated using diffusion tensor imaging. *AJNR Am J Neuroradiol*. 2008;
- 500 42. Yu Q, Qian W. Poor Spontaneous Recovery of Aphemia Accompanied by Damage to the Anterior Segment of the
501 Left Arcuate Fasciculus: A Case Report. *Brain Sci*. 2022 Sep 16;12(9):1253.
- 502 43. Ivanova MV, Zhong A, Turken A, Baldo JV, Dronkers NF. Functional Contributions of the Arcuate Fasciculus to
503 Language Processing. *Front Hum Neurosci*. 2021 Jun 25;15:672665.
- 504 44. Schevenels K, Gerrits R, Lemmens R, De Smedt B, Zink I, Vandermosten M. Early white matter connectivity and
505 plasticity in post stroke aphasia recovery. *NeuroImage Clin*. 2022;36:103271.
- 506 45. Bae CR, Na Y, Cho M, Hwang YM, Tae WS, Pyun SB. Structural Changes in the Arcuate Fasciculus and Recovery of
507 Post-stroke Aphasia: A 6-Month Follow-up Study using Diffusion Tensor Imaging. *Neurorehabil Neural Repair*.
508 2022 Sep;36(9):633–44.
- 509 46. Forkel SJ, Catani M. Lesion mapping in acute stroke aphasia and its implications for recovery. *Neuropsychologia*.
510 2018;Jul1(115):88–100.
- 511 47. Marchina S, Zhu LL, Norton A, Zipse L, Wan CY, Schlaug G. Impairment of speech production predicted by lesion
512 load of the left arcuate fasciculus. *Stroke*. 2011;42(8).
- 513 48. Keser Z, Sebastian R, Hasan KM, Hillis AE. Right Hemispheric Homologous Language Pathways Negatively
514 Predicts Poststroke Naming Recovery. *Stroke*. 2020 Mar;51(3):1002–5.
- 515 49. Moulton E, Magno S, Valabregue R, Amor-Sahli M, Pires C, Lehericy S, et al. Acute Diffusivity Biomarkers for
516 Prediction of Motor and Language Outcome in Mild-to-Severe Stroke Patients. *Stroke*. 2019 Aug;50(8):2050–6.
- 517



Dorsal stream

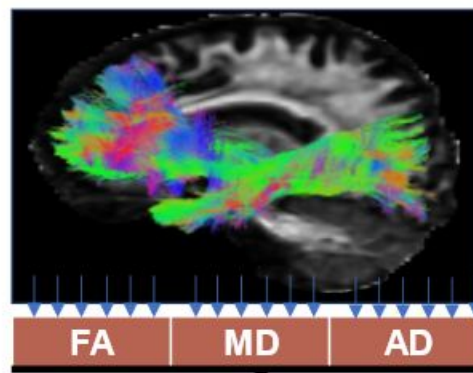
Arcuate Fasciculus (AF)
Superior Longitudinal Fasciculus (SLF)



Diffusion measures
obtained at acute phase

Ventral stream

Inferior Fronto-Occipital Fasciculus (IFOF)
Inferior Longitudinal Fasciculus (ILF)
Uncinate Fasciculus (UF)



1) BACKWARDS REGRESSIONS

Subacute language
measures
(CS_2 , ΔCS_{1-2} and
 $r\Delta CS_{1-2}$)

Chronic language
measures
(CS_3 , ΔCS_{2-3} and
 $r\Delta CS_{2-3}$)

Subacute language
measures
(CS_2 , ΔCS_{1-2} and
 $r\Delta CS_{1-2}$)

Chronic language
measures
(CS_3 , ΔCS_{2-3} and
 $r\Delta CS_{2-3}$)

Multicollinearity control :

- VIF > 5
- Condition index > 30

2) FILTERING COLLINEAR PAIRS

3) NESTED REGRESSIONS

Prediction of CS_2

Prediction of ΔCS_{1-2}

Prediction of $r\Delta CS_{1-2}$

Prediction of CS_3

Prediction of ΔCS_{2-3}

Prediction of $r\Delta CS_{2-3}$

Prediction of CS_3

Prediction of ΔCS_{2-3}

Prediction of $r\Delta CS_{2-3}$

Predictors in subacute prediction models:
 CS_1 + Age + Education + Corrected lesion size

Predictors in chronic prediction models (I):
 CS_1 + Age + Education + Corrected lesion size

Predictors in chronic prediction models (II):
 CS_2 + Age + Education + Corrected lesion size

4) MEDIATION ANALYSIS

