

1 **White matter plasticity during second language learning within and across hemispheres**

2

3 Xuehu Wei^{1*}, Thomas C. Gunter¹, Helyne Adamson¹, Matthias Schwendemann¹, Angela D.
4 Friederici¹, Tomás Goucha¹, and Alfred Anwander¹

5 ¹Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain
6 Sciences, Leipzig, Germany

7

8 *Corresponding author

9 Email: xuehuwei@cbs.mpg.de

10

11 Author Contributions: XW, AA, TG, and ADF designed the analysis protocol; TG, AA, MS,
12 and ADF conceived the experiment; TG, AA, MS collected the data, XW, HA, and AA
13 analyzed the data; XW, TCG, ADF, and AA wrote the paper; and XW, TCG, TG, ADF, and
14 AA reviewed the paper.

15

16

17

18 **Competing Interest Statement:** The authors declare no conflict of interest.

19 **Classification:** Biological Sciences: Neuroscience. Social Sciences: Psychological and
20 Cognitive Sciences

21

22 **Keywords:** Second language learning, Structural connectivity, Neuroplasticity

23

24

25 **Abstract**

26 Adult second language (L2) learning is a challenging enterprise inducing neuroplastic changes
27 in the human brain. However, it remains unclear how the structural language connectome and
28 its subnetworks change during adult L2-learning. The current study investigated longitudinal
29 changes in white matter (WM) language networks in each hemisphere, as well as their
30 interconnection, in a large group of Arabic-speaking adults who learned German intensively for
31 six months. We found a significant increase in WM-connectivity within bilateral temporal-
32 parietal semantic and phonological subnetworks and right temporal-frontal pathways mainly in
33 the second half of the learning period. At the same time, WM-connectivity between the two
34 hemispheres decreased significantly. Crucially, these changes in WM-connectivity are
35 correlated with L2 performance. The observed changes in subnetworks of the two hemispheres
36 suggest a network reconfiguration due to lexical learning. The reduced interhemispheric
37 connectivity may indicate a key role of the corpus callosum in L2-learning by reducing the
38 inhibition of the language-dominant left hemisphere. Our study highlights the dynamic changes
39 within and across hemispheres in adult language-related networks driven by L2 learning.

40

41 **Significance**

42

43 The neuroplastic changes induced by learning a second language (L2) in adulthood open
44 up new perspectives for understanding brain function. The current study shows structural
45 changes in the language network of Arabic native speakers who learned German intensively in
46 two phases of three months each. We found a marked change in the left-hemispheric lexical-
47 semantic language system and the right fronto-temporal pathway, accompanied by a decrease
48 in white matter connectivity in the corpus callosum during L2 learning, which occurred mainly
49 in the second period of L2 acquisition. The reduced interhemispheric connectivity suggests that
50 the inhibitory role of the corpus callosum, relevant for native language processing, is reduced
51 by L2 learning. Our findings demonstrate a clear experience-dependent structural plasticity in
52 the human brain during L2 learning.

53

54

55 **Introduction**

56 Cognitive functions develop in parallel with the plastic adaptation of the brain (1–6). This
57 suggests that the gray and white matter of the brain is altered by the acquisition of new skills
58 and thus is modulated by lifelong experiences, such as the acquired native language (7). Second
59 language (L2) learning in adulthood is a complex task that requires the adaptation of multiple
60 brain systems related to a wide range of novel tasks to be mastered. To date, changes associated
61 with L2 learning were reported to extend beyond the brain regions of the native language
62 network in the left-hemisphere (8–10), with additional involvement of the right-hemisphere (11,
63 12), as well as plasticity in the white matter connections between the two hemispheres (13, 14).
64 How these changes in the gray and white matter might develop during L2 learning is described
65 in a model (15) called the Dynamic Restructuring Model (DRM).

66 The DRM-model postulates three distinct phases of structural adaptation that depend on
67 the quantity and quality of the language learning and language switching experience. In the
68 earliest phase, L2 learning leads to changes in the gray matter areas that support the processing
69 of the new language. Next, in the intermediate consolidation phase, the white matter pathways
70 connecting the language processing areas show a structural modulation. Finally, in the peak
71 efficiency phase, the model predicts further changes in brain structure, including increased
72 frontal white matter connectivity, leading to highly efficient L2 processing and language
73 switching performance. However, longitudinal studies of white matter changes in L2 learning
74 in large samples of adults who have achieved proficiency beyond the beginner level are still
75 lacking. The present study aims to investigate different phases of longitudinal white matter
76 changes within each hemisphere and across hemispheres, to clarify the role of the corpus
77 callosum, and to describe the brain mechanisms involved in L2 learning.

78 L2 learning comprises the acquisition of a new vocabulary, which includes learning novel
79 phonemes, phonetic categories as well as word meanings, in addition to a new grammar. At the
80 behavioral level, it has been previously reported that lexical-semantic processing of newly
81 learned words and simple grammar is relatively easy to acquire, and native-like performance
82 can be achieved in L2 learners (16). In contrast, it is more difficult for late L2 learners to
83 perform real-time syntactic analysis, and they do not achieve automatic, highly proficient
84 syntax processing until a late stage of learning (12, 16).

85 At the neurofunctional level, brain imaging studies have shown that low proficient L2
86 learners have less overlap in brain activation between first and second language processing than
87 high proficient L2 learners (17) and recruit additional brain areas in the right hemisphere (18).
88 These brain areas may support language proficiency by effectively handling word retrieval (11).

89 Comparing first and second language brain activation in the lexical-semantic domain was found
90 to depend on the learners' performance, but differences in the grammatical domain on the age
91 of L2 acquisition (19).

92 Studies focusing on the neuroplasticity of the language system as a function of L2 learning
93 (9, 10) have reported changes in the gray matter of the bilateral Inferior Frontal Gyrus (IFG),
94 Inferior Parietal Lobe (IPL), and anterior and posterior Temporal Lobe (TL) (13, 20–23). In
95 particular, they include cortical gray matter changes in the bilateral TL and IPL, related to
96 phonological and lexical-semantic memory systems that are crucial for the acquisition of the
97 new vocabulary (24–28). Additionally, since languages differ in their syntactic and
98 morphological rules, successful L2 acquisition also depends on the brain's adaptation to
99 grammatical processing (29, 30). In native language processing, semantics and grammatical
100 rules are processed in a left-lateralized network including inferior frontal and temporal-parietal
101 regions, which are connected via dorsal and ventral white matter pathways (31). L2 acquisition
102 during adulthood requires neural adaptations that reach beyond the classical language network,
103 involving the right hemisphere (8–10), playing an essential role in the early learning phases
104 when L2 processing is not yet fully automatized (24, 32).

105 In addition to the reported changes in gray matter, plasticity of the white matter language
106 pathways in L2 learning (9) has also been suggested in previous cross-sectional studies
107 comparing bilinguals and monolinguals (20, 23, 33), as well as in some longitudinal language
108 learning studies (13, 22). These studies have shown an association between L2 acquisition and
109 local changes in white matter parameters which were located in the bilateral Inferior Fronto-
110 Occipital Fascicle (IFOF), the Superior Longitudinal Fascicle (SLF), the Arcuate Fascicle (AF),
111 the Uncinate Fascicle (UF) and the Corpus Callosum (CC) (13, 20, 22, 23, 33) which might be
112 related to alterations in myelination or axonal characteristics (2, 34). White matter plasticity has
113 been reported in relation to different aspects (i.e., novel speech sounds, vocabulary, grammar,
114 etc.) of L2 acquisition (20, 24, 35) and respective variations across the phases of language
115 learning (9, 10, 36).

116 Although it is widely accepted that language processing is dominated by the left
117 hemisphere (31, 37), increasing evidence suggests that the right hemisphere is highly involved
118 in L2 learning (10, 38, 39), including dynamic changes in lateralization across phases of
119 language learning (40). In addition, differences in the CC have been reported between bilingual
120 and monolingual participants, suggesting its involvement in L2 learning (9, 14). The CC is the
121 structural bridge that allows the interaction between the hemispheres (41, 42). However, its role
122 in the acquisition and use of a second language remains unclear.

123 There are two competing theories for the general role of the CC in interhemispheric
124 interaction. One argues for the inhibition of the activation in the other hemisphere and the other
125 suggests excitatory mechanisms (for reviews see (43, 44)). The interhemispheric inhibition
126 theory proposes that an area can reduce the activity in homologous contralateral areas via the
127 CC to allow for fast and automatic processing within each hemisphere and leading to functional
128 hemispheric specialization. In contrast, the excitatory theory suggests that activation in one
129 hemisphere facilitates the activation of homolog areas, increasing information exchange
130 between the hemispheres. It is still an open question whether, during initial L2 learning, the
131 role of the CC is mainly excitatory or inhibitory. However, it is well established that first
132 language (L1) processing is strongly left lateralized and the CC establishes a strong inhibition
133 from the dominant left hemisphere on the right hemisphere. In contrast, early phases of L2
134 learning involve the right hemisphere possibly due to weakened inhibition in early phases of
135 L2 learning, hence allowing the engagement of the right homologs of the language areas.
136 Accordingly, a decrease in CC-connectivity would follow a reduction of inhibitory explanation
137 suggesting that this mechanism might play a role in L2 learning.

138 Here, we provide empirical data of longitudinal white matter changes as a function of L2
139 learning in two successive phases. Based on the DRM-model, we hypothesize that L2 learning-
140 induced changes in structural connectivity will occur mainly after an initial beginner phase of
141 learning, during which white matter changes are expected to be limited. In a second,
142 intermediate consolidation phase, we expect changes in the structural connectivity of the
143 language network. This learning phase involves both semantic processing of new vocabulary,
144 and on the other hand, local syntactic processing based on lexical word category and semantic
145 information. We hypothesize that these plastic changes take place primarily in the lexical-
146 semantic system of both hemispheres and that such changes are related to the improvement in
147 L2 performance. Furthermore, we hypothesize that L2 learning will lead to a significant change
148 in transcallosal connectivity, supporting the role of the CC in interhemispheric communication.

149 To test these hypotheses, we recruited a large group of young, healthy Arabic native-
150 speaking participants for an intensive German language course over six months to reach an
151 intermediate proficiency (B1) level. The course consisted of an initial beginner phase of three
152 months and a consolidation phase of the same duration. After three and six months, the
153 participants took a standardized German language test that assessed L2 comprehension and
154 production. At the beginning of the course and after each learning period, we acquired
155 longitudinal high-resolution diffusion MR images and computed the white matter structural
156 connectivity network in each participant. This structural network included the intrahemispheric

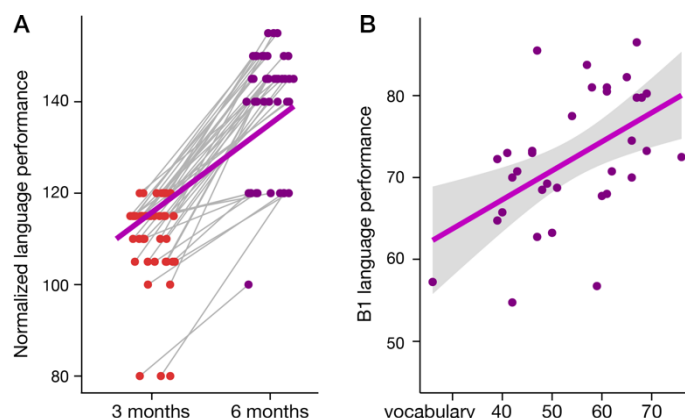
157 connections between the language processing areas in the left hemisphere, and between their
158 right hemisphere homologs as well as the callosal connections of these cortical areas in both
159 hemispheres. Then, we compared the network properties between the different time points to
160 identify changes in specific pathways and subnetworks. To test all connections for longitudinal
161 changes in connectivity in an unbiased manner without introducing strong *a priori* hypotheses
162 into the analysis, we used the recently proposed network-based statistics (NBS) (45) and a
163 mixed-effects model (46). Changes in structural connectivity were then related to improvements
164 in language tests to demonstrate a direct functional relevance of the detected changes in the
165 language network.

166

167 Results

168 **Improvement of L2 performance.** The L2 performance after three months (59 participants)
169 and after six months of learning (51 participants) was measured with standardized tests for
170 German as an L2. The results of both tests were normalized to a common scale following the
171 scaling method proposed in the Cambridge English Scale. Linear mixed-effects (LME) models
172 were used in MATLAB to analyze behavioral improvement during learning, with time points
173 modeled as a fixed effect (see Methods). The data showed a significant improvement in L2
174 performance between the two time points of the German tests ($t=17.92$, $p<0.0001$, see [Figure](#)
175 [1A](#)). After six months of learning, 41 participants took an additional vocabulary test.
176 Correlation analysis between L2 vocabulary and the B1 language test showed that individuals
177 with richer L2 vocabulary had higher overall language proficiency ($r=0.509$, $p=0.002$, see
178 [Figure 1B](#)).

179



180

181 **Figure 1. L2 improvement during learning.** (A) Longitudinal changes of the
182 normalized language performance after three and six months of L2 learning. (B)

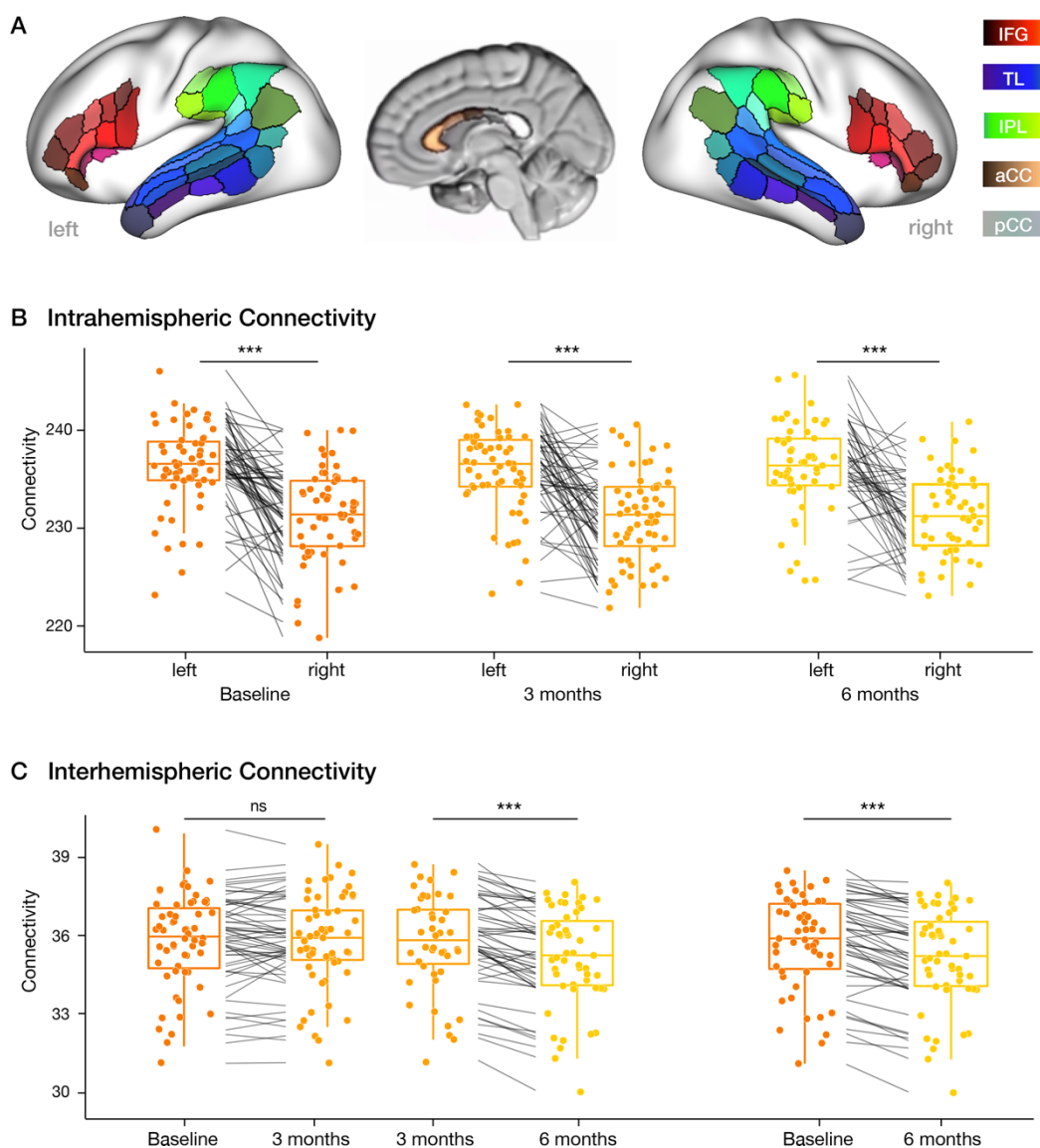
183 Correlation of L2 vocabulary score and overall language performance (B1 test) after
184 six months of L2 learning.

185

186 **Lateralization and longitudinal changes of the intra- and interhemispheric connectivity.**

187 The initial lateralization test showed that the global intra-hemispheric connectivity in the
188 language network is stronger in the left hemisphere than in the right hemisphere for each of the
189 three measurement time points (baseline: left > right, $t = 8.17$, $p < 0.0001$; 3 months: left > right,
190 $t = 6.71$, $p < 0.0001$; 6 months: left > right, $t = 6.56$, $p < 0.0001$; see [Figure 2B](#)). The longitudinal
191 statistical analysis was then performed separately for the total intra-hemispheric connectivity
192 in each side and the interhemispheric connectivity using an LME model with time points as a
193 fixed effect. The result showed a significant dynamic decrease in interhemispheric connectivity
194 during learning, specifically with an effect in the second half of the learning period (baseline-3
195 months: $t = -0.56$, $p = 0.57$ (n.s.); 3 months – 6 months: $t = -7.33$, $p < 0.0001$, baseline – 6
196 months: $t = -8.97$, $p < 0.0001$; see [Figure 2C](#)). However, we did not observe any significant
197 changes in the longitudinal analysis of intra-hemispheric connectivity within the language
198 network in each hemisphere or the lateralization index of the connectivity within the language
199 network.

200

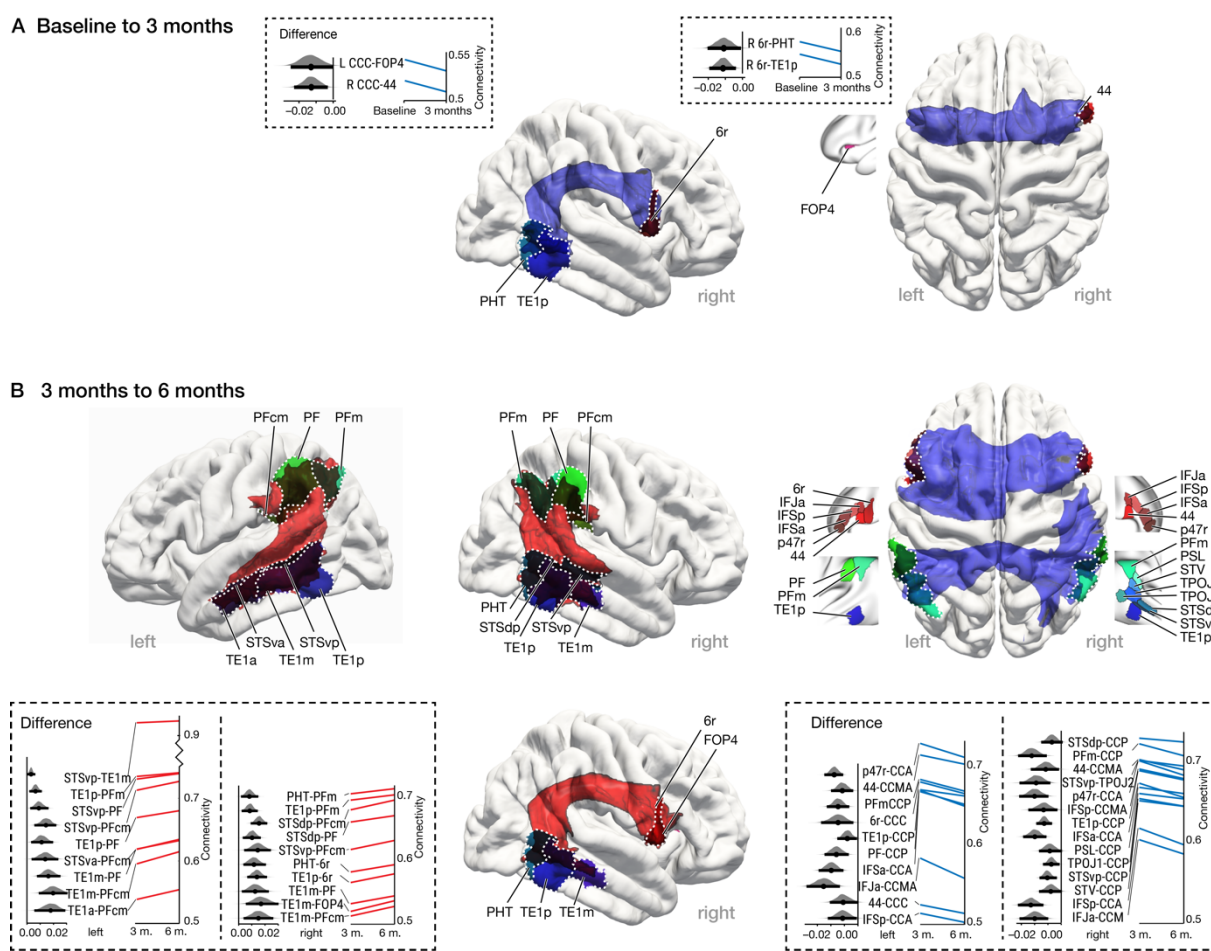


201

202 **Figure 2. Longitudinal changes of the intra- and interhemispheric connectivity**
203 **in the language network.** (A) Areas in the language network of the left and right
204 hemisphere in the inferior frontal gyrus (IFG), the inferior parietal lobe (IPL) the
205 superior and middle temporal lobe (TL), and the anterior and posterior corpus
206 callosum (aCC, pCC). (B) Intra-hemispheric connectivity at each time point during
207 L2 learning shows significant left lateralization of the language network. (C)
208 Longitudinal changes in interhemispheric connectivity show a significant decrease
209 in the second learning period (middle) and over the full 6 months (right). The
210 boxplots show the median, quartiles, 1.5* interquartile range, and all individual data
211 points. (***) $p < 0.0001$.

212

213 **Plasticity of the structural language subnetworks across different learning periods.** The
214 Network-Based R-Statistics (NBS) LME models (p-threshold = 0.01, K = 5000 permutations)
215 revealed a complex reorganization of multiple subnetworks during L2 learning, including
216 connections between all subregions in the bilateral temporal lobe (TL), inferior parietal lobe
217 (IPL), and right inferior frontal gyrus (IFG, $p < 0.01$, NBS corrected, [Supplementary Figure](#)
218 [S2](#)). To shed light on the temporal properties of network changes, post hoc LME analyses
219 between adjacent time points allowed us to identify specific effects of each connectivity for the
220 early and later learning period ($p < 0.05$). During the first three months of learning, there was a
221 significant decrease in connectivity for only a few connections belonging to three subnetworks.
222 These subnetworks consisted of the interhemispheric connections of subregions of the IFG and
223 parts of the right arcuate fascicle (AF) connecting the posterior IFG and the posterior middle
224 and inferior temporal gyrus ([Figure 3A](#)). However, in the second learning period (from three to
225 six months), the statistical analysis revealed an increase in connectivity in three subnetworks,
226 including the bilateral parietal-temporal system as well as the right AF (see [Figure 3B](#), left).
227 Interestingly, the frontal and temporal-parietal interhemispheric networks showed decreased
228 connectivity in this second period (see [Figure 3B](#), right). The figure shows the mean changes
229 of all individual connections within each subnetwork and the distribution of the changes.
230 Individual data for each participant and each connection are shown in [Supplementary Figure](#)
231 [S4](#).
232



233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

Figure 3. Subnetworks with longitudinally increased and decreased connectivity in the two learning periods. (A) First learning period: Decreasing connectivity (blue) in three small subnetworks including anterior transcallosal connections as well as the right AF. (B) Second learning period: Increasing connectivity (red) in three large subnetworks connecting the posterior temporal and the inferior parietal regions in both hemispheres along with the right AF (left). Decreasing connectivity (blue) in the anterior and posterior transcallosal subnetworks (right, all $p < 0.05$ NBS corrected). The brain figure shows the group averaged probabilistic tractography of the subnetworks with increased (red) and decreased (blue) connectivity together with the corresponding brain regions. The figure in the box shows the effect size and change trend of each connection.

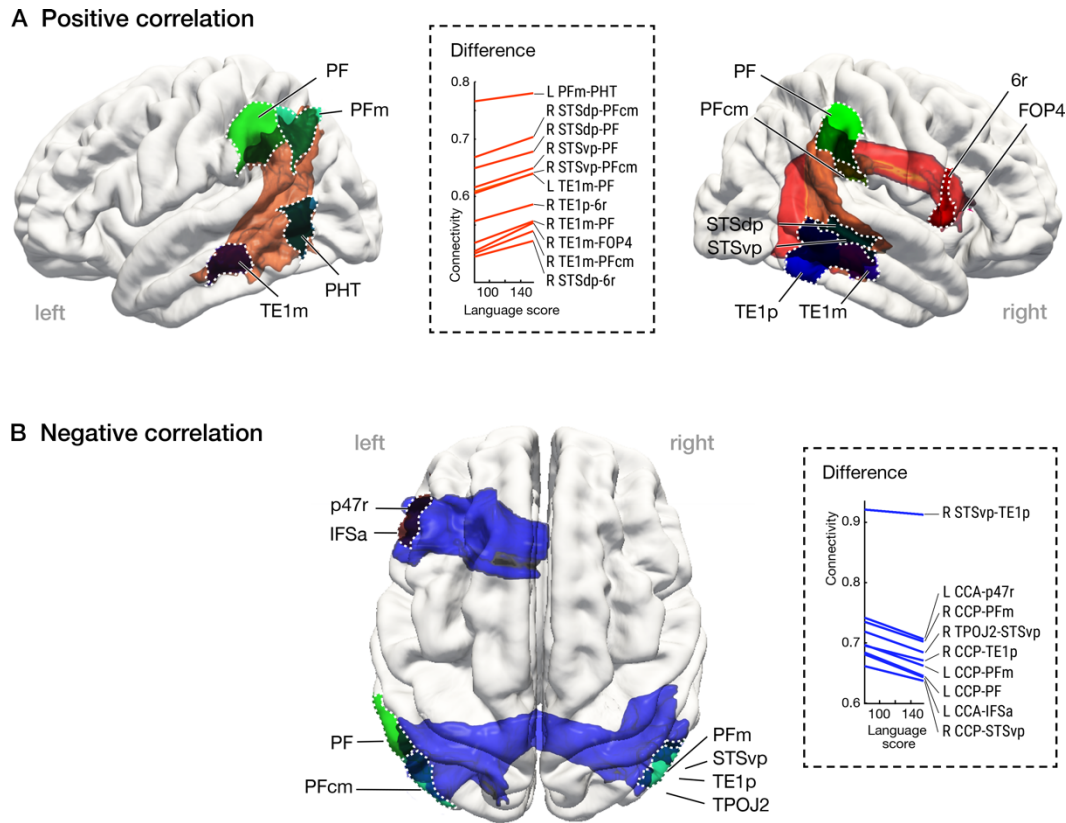
Relationship between L2 proficiency and connectivity changes in the language network.

To test the relationship between brain network plasticity and L2 performance increase over the different learning periods, we also used NBS with LME models. The initial behavioral analysis revealed that all participants showed an improvement in their L2 scores between three and six

250 months of learning (Figure 1). This monotonic increase in performance allowed us to use a
251 more parsimonious LME model that included only the L2 score and did not require modeling
252 time as an additional separate factor. This NBS LME model allowed us to test for longitudinal
253 correlations between the structural network characteristics and L2 performance for each
254 participant after three and six months of learning. Figure 4 shows the brain subnetworks that
255 show a significant linear relationship between the L2 score and the brain connectivity at three
256 and six months ($p < 0.01$, NBS corrected). The NBS LME analysis showed that the improvement
257 in L2 proficiency was correlated with increased connectivity in subnetworks connecting the
258 posterior temporal and the inferior parietal lobes in both hemispheres as well as in the right
259 arcuate fascicle (AF) (see Figure 4A). Additionally, a negative correlation between connectivity
260 changes and the L2 score was found in the anterior and posterior interhemispheric connections
261 ($p < 0.01$, NBS corrected, Figure 4B). In the frontal lobe, only the left subnetwork of the
262 transcallosal connections showed a significant correlation. The figure shows the regression
263 lines of the correlation for all individual connections within each subnetwork. The individual
264 data for each participant and each connection are shown in Supplementary Figure S5.

265 Another measure of L2 learning success relates to L2 vocabulary size. In a *post hoc* analysis,
266 we examined whether participants with a large L2 vocabulary showed different changes in
267 white matter connectivity compared to those with a smaller L2 vocabulary. The group was
268 divided based based on their productive vocabulary in the written text of the B1 test. After
269 splitting into two groups, only the group with higher vocabulary scores showed a positive
270 correlation between increased L2 scores and connectivity changes in the right temporal-parietal
271 and AF subnetwork and a negative correlation with interhemispheric connectivity ($p < 0.01$,
272 NBS corrected, Supplementary Figure S3). This suggests an important role of the right
273 hemisphere for successful L2 learning.

274



275

276

277

278

279

280

281

282

283

284

285

286 Discussion

287

288

289

290

291

292

293

Figure 4. Relating changes in connectivity and the L2 proficiency from three to six months of L2 learning. (A) Positive correlation between L2 performance and connectivity changes in the left and right temporal-parietal network and the AF (red). (B) Negative correlation in the anterior (left) and posterior transcallosal network (blue, all $p < 0.05$ NBS corrected). The plots display the regression lines for all connections in all networks. In the brain images, the colored mean tractography shows the correlated subnetworks together with the corresponding brain regions. The figure in the box shows the correlation trend between each connectivity and L2 proficiency.

294 bilateral temporal-parietal system and the right arcuate fascicle (AF) showed increased
295 connectivity. Additionally, the interhemispheric connectivity across the corpus callosum (CC)
296 was reduced during this learning period. Most importantly, the brain changes in the late learning
297 period correlated with the increases in L2 proficiency.

298 Our findings provide empirical evidence for the time course and location of white matter
299 changes in the consolidation phase as suggested by the Dynamic Restructuring Model (DRM)
300 of second language acquisition (15). This model suggests that plasticity in the white matter
301 language network will emerge in a second phase of L2 acquisition, allowing for more efficient
302 interaction between the different language areas within each hemisphere. This is indeed what
303 we found in the present study.

304

305 **Dynamic intra-hemispheric white matter changes underlying L2 performance.**

306 The white matter network is the structural basis for neuronal communication between brain
307 areas and its plasticity is crucial for learning new skills (2). Distinct subnetworks within the
308 language system are specialized for different domains, and the corresponding connections are
309 modulated by their usage. Therefore, we expected changes in the subnetworks reflecting
310 specific tasks to be mastered in the newly learned language.

311 *Semantic and phonological processes:* The analysis of the language test showed a
312 significant improvement in L2 from three to six months. L2 performance at six months
313 correlated with the results of an independent L2 vocabulary test (47). From a neuroplasticity
314 perspective, it is important to note that we found a significant change in the structural
315 connectivity over this period in bilateral temporal-parietal subnetworks and right temporal-
316 frontal connections. Crucially, these longitudinal changes in connectivity were found in
317 subnetworks that were very similar to those that showed plastic changes correlated with L2
318 proficiency in our study. These subnetworks form the structural basis for lexical-semantic and
319 phonological processing (31, 48, 49) and have previously been related to L2 vocabulary
320 learning (24, 25, 50). In this regard, it is interesting to look at subgroups of participants with
321 higher and lower L2 vocabulary scores. Our additional correlational analysis in these subgroups
322 revealed a significant positive correlation between L2 improvement and changes in connectivity
323 in the right fronto-temporal subnetwork as part of the AF in the second learning phase only in
324 the group with a larger L2 lexicon compared to the other group (see [Supplementary Figure S3](#)).
325 This result suggests that during this phase of L2 learning, the changes in the language network
326 are related to the consolidation of lexical processing and highlights the importance of the right
327 hemisphere for L2 acquisition.

328 Successful language learning depends on phonological discrimination during perception
329 and phonological selection during production to decode speech sounds and associate them with
330 the meaning of new words. In the neural language network, the bilateral inferior parietal lobe
331 (IPL) and the superior temporal gyrus (STG) are involved in phonological storage and word
332 decoding, and the middle temporal gyrus (MTG) is an integral region engaging in lexical-
333 semantic access (31, 49, 51) Functional imaging studies of L2 learning (30) suggest a stronger
334 functional activity and connectivity of the right IPL and STG during early learning phases, and
335 phonological processing of L2 words (52, 53). Structural imaging studies demonstrate that these
336 regions show changes in gray matter morphology during L2 learning that are related to L2
337 vocabulary acquisition and L2 proficiency (25, 26, 50, 52–55).

338 In addition to white matter effects in temporal-parietal connections, we also found
339 increased connectivity in the right hemispheric temporal-frontal subnetwork as part of the AF
340 which corresponds to the right hemispheric equivalent of the dorsal language network. This
341 finding is consistent with previous studies highlighting the importance of the right hemisphere
342 for L2 lexical-semantic and phonological processing during the initial and intermediate phases
343 of adult L2 acquisition (10, 24). In addition, a higher involvement of the right prefrontal cortex
344 (56) during L2 processing may be related to more working memory and attentional processes
345 in L2 (57). Taken together, our findings strongly suggest that efficient L2 learning, especially
346 vocabulary acquisition in adults, involves the right fronto-parieto-temporal network.

347

348 **Longitudinal decrease in transcallosal interhemispheric connectivity.**

349 The observed network strength in both hemispheres showed that the language network is
350 lateralized to the left at all measured time points during L2 learning. This is in line with the
351 widely accepted model that the language network is dominated by the left hemisphere (31). In
352 previous studies (10, 32), and the present data, there is evidence of increased right hemisphere
353 involvement during L2 learning, as reflected by strong changes in white matter connectivity in
354 the right hemisphere. These changes might be directly related to the reduced transcallosal
355 connectivity, allowing for additional L2 processing to occur in the right hemisphere (43).
356 Indeed, we found a significant decrease in the interhemispheric connectivity in the anterior and
357 posterior CC during L2 learning. This reduction correlated with the increase in L2 performance
358 in the second learning phase. The current data provide a comprehensive demonstration of the
359 role of the CC in L2 learning. In native language processing, the dominant left hemisphere
360 exerts an inhibitory influence on the non-dominant right hemisphere via the CC (44). However,
361 during the initial and intermediate phases of L2 learning, a highly involved right hemisphere

362 language network is required to build up the L2 lexicon. This would explain why successful L2
363 acquisition is accompanied by a decrease in transcallosal connectivity. This reduces the
364 inhibition of the language-dominant left hemisphere on the corresponding regions in the right
365 hemisphere, allowing increased processing and connectivity to occur in the right half of the
366 brain. However, this fundamentally new finding needs to be further explored and supported by
367 additional data.

368

369 **Conclusions**

370 Our study showed that L2 learning in adults leads to dynamic changes in brain connectivity
371 within and across hemispheres. The experimental evidence suggests that plastic changes in the
372 white matter system occur mainly after an initial period of learning. In this phase, the adaptation
373 of the language network appears to be focused on the lexical-semantic system, particularly in
374 the temporal and temporal-parietal regions, with increased connectivity in each of the two
375 hemispheres and strong involvement of the right side. At the same time, L2 learning leads to
376 reduced connectivity between the hemispheres, which could result in reduced inhibition of the
377 dominant left hemisphere on the right, temporarily freeing up resources in the right half of the
378 brain for efficient learning of new L2 linguistic features.

379

380 **Materials and Methods**

381 **Participants.** Eighty-four young, healthy right-handed Arabic native speakers were
382 recruited for an intensive German course (5 h/day, 5 days/week) over six months to reach the
383 threshold level B1 according to the Common European Framework of Reference for Languages
384 (CEFR, (58)). The course was divided into two learning phases of three months each. During
385 the two learning periods, some participants left the language course for personal reasons and
386 were not included in the corresponding analysis. Fifty-nine participants completed the first
387 learning phase (mean age, 24.4 ± 4.5 (SD) years, 51 male) and 51 participants completed the
388 two learning phases (mean age, 24.7 ± 4.6 (SD) years, 43 male). After each learning phase,
389 participants took a 90-minute standardized second language proficiency test (A1 and B1 tests
390 of the Goethe Institute). After six months of learning, an additional L2 Vocabulary Size Test
391 (VST) was taken by a subgroup of 41 participants (35 male). All participants were immersed in
392 the second language environment and lived in Germany during the course. All participants were
393 native speakers of the Levantine dialect of Arabic and of normal intelligence (non-verbal
394 Raven's matrix test (59), score 50.4 ± 6.7 , ranging around the upper 90 percentile of the
395 reference population, subgroup N=32) and spoke only one native language. All participants

396 were recruited in Leipzig and arrived in Germany 6-8 months before the start of the study. They
397 were settled in Leipzig for a long-term stay and were highly motivated to learn German and
398 integrate into the academic system. An initial German test revealed that the entire group had
399 no-to-minimal knowledge of German, well below beginner level A1. Structural and high-
400 angular and spatial resolution diffusion MRI data were acquired from each participant on a
401 Siemens 3T Prisma MRI scanner at baseline and after three and six months of learning. Details
402 of the learning procedure and MRI acquisition can be found in the [Supplementary Materials](#).
403 The study was approved by the Ethics Committee of the University of Leipzig, and all
404 participants gave written informed consent in their native language.

405
406 **Structural language connectome.** We used probabilistic diffusion MRI tractography to
407 compute the white matter network between the language processing regions ([Figure 2A](#)) in each
408 participant and time point (baseline, three and six months of learning). The analysis followed
409 the previously established method (7). Cortical seed and target areas were defined using the
410 Human Connectome Project (HCP) fine-grained atlas in addition to a subdivision of the corpus
411 callosum (CC) atlas (60, 61). The core regions of the language network, as defined previously
412 (31), included the dorsal and ventral pathways in both hemispheres between subregions in the
413 bilateral inferior frontal gyrus (IFG), superior temporal gyrus (STG), middle temporal gyrus
414 (MTG), and inferior parietal lobe (IPL). To account for interhemispheric connections, we
415 included white matter regions in the medial cross-section of the CC resulting in 33 cortical and
416 5 CC regions per hemisphere (see [Figure 2A](#) and [Supplementary Materials, Table S2](#)). The CC
417 is a bottleneck for estimating interhemispheric connections, and direct one-to-one connectivity
418 between cortical areas in both hemispheres cannot be robustly estimated by tractography.
419 Therefore, we computed probabilistic tractography between cortical and CC regions as a robust
420 approximation of the interhemispheric connectivity. To remove false-positive connections (62),
421 we retained the 30% strongest connections for the network analysis. Details of the connectivity
422 analysis can be found in the [Supplementary Materials](#).

423
424 **Statistical analyses.** To estimate longitudinal changes in the behavioral learning progress,
425 a linear mixed-effects model (LME) (i.e., $y \sim \text{time} + (1 \mid \text{participant})$; y represents L2
426 proficiency) with time points as the fixed effect and participant as the random effect was applied
427 to the scaled language test scores obtained after three and six months of learning to analyze
428 changes across learning periods. In addition, a correlation analysis was performed between the

429 L2 vocabulary level after six months and the B1 language test scores to analyze how strongly
430 the composed language test is related to vocabulary knowledge at this stage in this group.

431 To assess the relationship between L2 learning and plasticity in the white matter language
432 network, we first tested the changes in the overall network strength within and between brain
433 hemispheres. We first measured the network strength in each hemisphere (sum of all
434 connectivity values between the cortical regions) and compared this parameter between
435 hemispheres at each time point (baseline, three and six months of learning) to test for
436 lateralization of the language network using a paired t-test. Next, we used a separate LME
437 models (i.e., $y \sim \text{time} + (1 | \text{participant})$; y represents connectivity) with the three measurement
438 time points as a fixed effect to test longitudinal learning-induced changes in interhemispheric
439 connectivity as well as intra-hemispheric changes in the language network within each
440 hemisphere (left and right) and the lateralization index. The interhemispheric network strength
441 is the sum of the weighted connections between all cortical language regions and the CC,
442 representing the connections crossing to the other hemisphere. In all analysis steps discussed
443 so far, the LME models used time as a fixed effect and participant as a random effect, and the
444 statistical tests were performed in MATLAB.

445 To localize subnetworks showing longitudinal changes within the language network across
446 all time points, the recently proposed network-based R-statistics (NBS) (45) for LME models
447 (46) with time points as a fixed effect and participant as a random effect was used. In the first
448 step, the connectivity measures at different time points were modeled as fixed effects and
449 participants were entered as random effects. This allowed us to analyze the longitudinal change
450 of the structural connectome across the three different time points and account for individual
451 differences (i.e., $y \sim \text{time} + (1 | \text{participant})$; y represents connectivity). Then, posthoc statistics
452 were used to identify subnetworks with significant changes between each pair of measurement
453 points and to determine in which of the two language learning phases white matter changes
454 occurred. Finally, to test whether changes in connectivity were related to individual L2
455 performance and to localize such subnetworks, we employed a second type of model and
456 included L2 proficiency test scores as a fixed effect and participant as a random effect in the
457 NBS LME models, allowing us to account for the longitudinal interactions between brain
458 structural connectivity and L2 proficiency (i.e., $y \sim \text{score} + (1 | \text{participant})$; y represents
459 connectivity). Since test scores could only be acquired after three and six months of learning,
460 these two time points were considered in this analysis. To visualize the statistically identified
461 subnetworks, additionally, probabilistic tractography was computed between the regions
462 belonging to these subnetworks. The individual pathway maps were normalized, averaged, and

463 visualized together with the regions of the subnetwork. This allowed to show the white matter
464 pathways belonging to the identified subnetworks.

465

466 **Acknowledgments.** We would like to thank S. Brogatio of the former Leipzig Refugee Council
467 Leipzig for her help in recruiting the participants as well as all participants, teachers, and student
468 helpers for their contributions. We thank Prof. C. Fandrych, Herder Institute, Univ. Leipzig for
469 his support in the conception and design of the language courses, and Prof. E. Tschirner, ITT,
470 Univ. Leipzig, for his support in the language tests, especially in the application of the
471 vocabulary tests. Finally, we would like to thank M. Lišaník for his contribution to data
472 acquisition, participant testing, and study conception. This work was supported by the SPP2041
473 program "Computational Connectomics" of the German Research Foundation (DFG), grant
474 number 347141397 (AN 1156/1-1, FR 519/22-1).

475

476 **References**

477

478

- 479 1. R. D. Fields, White matter in learning, cognition and psychiatric disorders. *Trends*
480 *Neurosci.* **31**, 361–370 (2008).
- 481 2. R. J. Zatorre, R. D. Fields, H. Johansen-Berg, Plasticity in gray and white:
482 neuroimaging changes in brain structure during learning. *Nat. Neurosci.* **15**, 528–536
483 (2012).
- 484 3. B. Draganski, G. Christian, B. Volker, S. Gerhard, B. Ulrich, M. Arne, Changes in
485 grey matter induced by training. *Nature.* **427**, 311–312 (2004).
- 486 4. E. A. Maguire, D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. J.
487 Frackowiak, C. D. Frith, Navigation-related structural change in the hippocampi of
488 taxi drivers. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4398–4403 (2000).
- 489 5. J. Scholz, M. C. Klein, T. E. J. Behrens, H. Johansen-berg, Training induces changes
490 in white matter architecture. *Nat. Neurosci.* **12**, 1370–1371 (2009).
- 491 6. M. Taubert, B. Draganski, A. Anwander, K. Muller, A. Horstmann, A. Villringer, P.
492 Ragert, Dynamic properties of human brain structure: learning-related changes in
493 cortical areas and associated fiber connections. *J. Neurosci.* **30**, 11670–11677 (2010).
- 494 7. X. Wei, H. Adamson, M. Schwendemann, T. Goucha, A. D. Friederici, A. Anwander,
495 Native language differences in the structural connectome of the human brain.
496 *Neuroimage.* **270**, 119955 (2023).

- 497 8. Z. Qi, M. Han, K. Garel, E. San Chen, J. D. E. Gabrieli, White-matter structure in the
498 right hemisphere predicts mandarin chinese learning success. *J. Neurolinguistics*. **33**,
499 14–28 (2015).
- 500 9. P. Li, J. Legault, K. A. Litcofsky, Neuroplasticity as a function of second language
501 learning: anatomical changes in the human brain. *Cortex*. **58**, 301–324 (2014).
- 502 10. Z. Qi, J. Legault, Neural hemispheric organization in successful adult language
503 learning: Is left always right? *Psychol. Learn. Motiv.* **72**, 119–163 (2020).
- 504 11. G. Videsott, B. Herrnberger, K. Hoenig, E. Schilly, J. Grothe, W. Wiater, M. Spitzer,
505 M. Kiefer, Speaking in multiple languages: neural correlates of language proficiency
506 in multilingual word production. *Brain Lang.* **113**, 103–112 (2010).
- 507 12. A. Hahne, A. D. Friederici, Processing a second language: Late learners’
508 comprehension mechanisms as revealed by event-related brain potentials. *Biling.*
509 *Lang. Cogn.* **4**, 123–141 (2001).
- 510 13. A. A. Schlegel, J. J. Rudelson, P. U. Tse, White matter structure changes as adults
511 learn a second language. *J. Cogn. Neurosci.* **24**, 1664–1670 (2012).
- 512 14. P. E. Coggins, T. J. Kennedy, T. A. Armstrong, Bilingual corpus callosum variability.
513 *Brain Lang.* **89**, 69–75 (2004).
- 514 15. C. Pliatsikas, Understanding structural plasticity in the bilingual brain: The Dynamic
515 Restructuring Model. *Biling. Lang. Cogn.* **23**, 459–471 (2020).
- 516 16. H. Clahsen, C. Felser, How native-like is non-native language processing? *Trends*
517 *Cogn. Sci.* **10**, 564–570 (2006).
- 518 17. S. A. Rüschemeyer, C. J. Fiebach, V. Kempe, A. D. Friederici, Processing lexical
519 semantic and syntactic information in first and second language: fMRI evidence from
520 German and Russian. *Hum. Brain. Mapp.* **25**, 266–286 (2005).
- 521 18. E. Higby, J. Kim, L. K. Obler, Multilingualism and the brain. *Annu. Rev. Appl.*
522 *Linguist.* **33**, 68–101 (2013).
- 523 19. I. Wartenburger, H. R. Heekeren, J. Abutalebi, S. F. Cappa, A. Villringer, D. Perani,
524 Early setting of grammatical processing in the bilingual brain. *Neuron*. **37**, 159–170
525 (2003).
- 526 20. P. K. Kuhl, J. Stevenson, N. M. Corrigan, J. J. F. van den Bosch, D. D. Can, T.
527 Richards, Neuroimaging of the bilingual brain: Structural brain correlates of listening
528 and speaking in a second language. *Brain Lang.* **162**, 1–9 (2016).
- 529 21. S. Murray Sherman, The thalamus is more than just a relay. *Curr. Opin. Neurobiol.*
530 **17**, 1–7 (2007).

- 531 22. P. C. Mamiya, T. L. Richards, B. P. Coe, E. E. Eichler, P. K. Kuhl, Brain white matter
532 structure and COMT gene are linked to second-language learning in adults. *Proc.*
533 *Natl. Acad. Sci. U. S. A.* **113**, 7249–7254 (2016).
- 534 23. C. Pliatsikas, E. Moschopoulou, J. D. Saddy, The effects of bilingualism on the white
535 matter structure of the brain. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1334–1337 (2015).
- 536 24. C. Hosoda, K. Tanaka, T. Nariyai, M. Honda, T. Hanakawa, Dynamic neural network
537 reorganization associated with second language vocabulary acquisition: a multimodal
538 imaging study. *J. Neurosci.* **33**, 13663–13672 (2013).
- 539 25. J. Legault, S. Y. Fang, Y. J. Lan, P. Li, Structural brain changes as a function of
540 second language vocabulary training: Effects of learning context. *Brain Cogn.* **134**,
541 90–102 (2019).
- 542 26. J. Mårtensson, J. Eriksson, N. C. Bodammer, M. Lindgren, M. Johansson, L. Nyberg,
543 M. Lövdén, Growth of language-related brain areas after foreign language learning.
544 *Neuroimage.* **63**, 240–244 (2012).
- 545 27. F. M. Richardson, M. S. C. Thomas, R. Filippi, H. Harth, C. J. Price, Contrasting
546 effects of vocabulary knowledge on temporal and parietal brain structure across
547 lifespan. *J. Cogn. Neurosci.* **22**, 943–954 (2010).
- 548 28. E. B. Barbeau, X. J. Chai, J. K. Chen, J. Soles, J. Berken, S. Baum, K. E. Watkins, D.
549 Klein, The role of the left inferior parietal lobule in second language learning: An
550 intensive language training fMRI study. *Neuropsychologia.* **98**, 169–176 (2017).
- 551 29. S. Caffarra, N. Molinaro, D. Davidson, M. Carreiras, Second language syntactic
552 processing revealed through event-related potentials: an empirical review. *Neurosci.*
553 *Biobehav. Rev.* **51**, 31–47 (2015).
- 554 30. K. M. Tagarelli, K. F. Shattuck, P. E. Turkeltaub, M. T. Ullman, Language learning in
555 the adult brain: A neuroanatomical meta-analysis of lexical and grammatical learning.
556 *Neuroimage.* **193**, 178–200 (2019).
- 557 31. A. D. Friederici, The brain basis of language processing: from structure to function.
558 *Physiol. Rev.* **91**, 1357–1392 (2011).
- 559 32. Z. Qi, M. Han, Y. Wang, C. de los Angeles, Q. Liu, K. Garel, E. S. Chen, S.
560 Whitfield-Gabrieli, J. D. E. Gabrieli, T. K. Perrachione, Speech processing and
561 plasticity in the right hemisphere predict variation in adult foreign language learning.
562 *Neuroimage.* **192**, 76–87 (2019).

- 563 33. E. Rossi, H. Cheng, J. F. Kroll, M. T. Diaz, S. D. Newman, Changes in white-matter
564 connectivity in late second language learners: Evidence from diffusion tensor
565 imaging. *Front. Psychol.* **8**, 2040 (2017).
- 566 34. W. Xin, J. R. Chan, Myelin plasticity: sculpting circuits in learning and memory. *Nat.*
567 *Rev. Neurosci.* **21**, 682–694 (2020).
- 568 35. V. DeLuca, K. Segaert, A. Mazaheri, A. Krott, Understanding bilingual brain function
569 and structure changes? U bet! A unified bilingual experience trajectory model. *J.*
570 *Neurolinguistics* **56**, 100930 (2020).
- 571 36. D. Perani, The neural basis of language talent in bilinguals. *Trends Cogn. Sci.* **9**, 211–
572 213 (2005).
- 573 37. A. D. Friederici, S. M. E. Gierhan, The language network. *Curr. Opin. Neurobiol.* **23**,
574 250–254 (2013).
- 575 38. H. Liu, F. Cao, L1 and L2 processing in the bilingual brain: A meta-analysis of
576 neuroimaging studies. *Brain Lang.* **159**, 60–73 (2016).
- 577 39. H. R. P. Park, G. Badzakova-Trajkov, K. E. Waldie, Language lateralisation in late
578 proficient bilinguals: a lexical decision fMRI study. *Neuropsychologia.* **50**, 688–695
579 (2012).
- 580 40. H. Xiang, T. M. van Leeuwen, D. Dediu, L. Roberts, D. G. Norris, P. Hagoort, L2-
581 proficiency-dependent laterality shift in structural connectivity of brain language
582 pathways. *Brain Connect.* **5**, 349–361 (2015).
- 583 41. A. D. Friederici, D. Y. von Cramon, S. A. Kotz, Role of the corpus callosum in
584 speech comprehension: interfacing syntax and prosody. *Neuron.* **53**, 135–145 (2007).
- 585 42. V. R. Karolis, M. Corbetta, M. Thiebaut de Schotten, The architecture of functional
586 lateralisation and its relationship to callosal connectivity in the human brain. *Nat.*
587 *Commun.* **10**, 1417 (2019).
- 588 43. J. S. Bloom, G. W. Hynd, The role of the corpus callosum in interhemispheric transfer
589 of information: excitation or inhibition? *Neuropsychol Rev.* **15**, 59–71 (2005).
- 590 44. L. J. van der Knaap, I. J. M. van der Ham, How does the corpus callosum mediate
591 interhemispheric transfer? A review. *Behav. Brain Res.* **223**, 211–221 (2011).
- 592 45. A. Zalesky, A. Fornito, E. T. Bullmore, Network-based statistic: Identifying
593 differences in brain networks. *Neuroimage.* **53**, 1197–1207 (2010).
- 594 46. Z. Gracia-Tabuenca, S. Alcauter, NBR: Network-based R-statistics for (unbalanced)
595 longitudinal samples. *bioRxiv.* **373019** (2020).

- 596 47. E. Tschirner, Examining the validity and reliability of the ITT vocabulary size tests.
597 Research Papers in Assessment. *Univ. Leipzig*. **3** (2021).
- 598 48. S. J. Forkel, E. Rogalski, N. Drossinos Sancho, L. D'Anna, P. Luque Laguna, J.
599 Sridhar, F. Dell'Acqua, S. Weintraub, C. Thompson, M. M. Mesulam, M. Catani,
600 Anatomical evidence of an indirect pathway for word repetition. *Neurology*. **94**,
601 e594–e606 (2020).
- 602 49. G. Hickok, D. Poeppel, The cortical organization of speech processing. *Nat. Rev.*
603 *Neurosci.* **8**, 393–402 (2007).
- 604 50. J. Legault, A. Grant, S.-Y. Fang, P. Li, A longitudinal investigation of structural brain
605 changes during second language learning. *Brain Lang.* **197**, 104661 (2019).
- 606 51. E. F. Lau, C. Phillips, D. Poeppel, A cortical network for semantics: (de)constructing
607 the N400. *Nat. Rev. Neurosci.* **9**, 920–933 (2008).
- 608 52. G. Raboyeau, K. Marcotte, D. Adrover-Roig, A. I. Ansaldo, Brain activation and
609 lexical learning: the impact of learning phase and word type. *Neuroimage*. **49**, 2850–
610 2861 (2010).
- 611 53. K. Veroude, D. G. Norris, E. Shumskaya, M. Gullberg, P. Indefrey, Functional
612 connectivity between brain regions involved in learning words of a new language.
613 *Brain Lang.* **113**, 21–27 (2010).
- 614 54. L. Ghazi Saidi, V. Perlberg, G. Marrelec, M. Pélégriani-Issac, H. Benali, A. I. Ansaldo,
615 *et al.*, Functional connectivity changes in second language vocabulary learning. *Brain*
616 *Lang.* **124**, 56–65 (2013).
- 617 55. A. M. Grant, S. Y. Fang, P. Li, Second language lexical development and cognitive
618 control: A longitudinal fMRI study. *Brain Lang.* **144**, 35–47 (2015).
- 619 56. M. Vigneau, V. Beaucousin, P. Y. Hervé, G. Jobard, L. Petit, F. Crivello, E. Mellet,
620 L. Zago, B. Mazoyer, N. Tzourio-Mazoyer, What is right-hemisphere contribution to
621 phonological, lexico-semantic, and sentence processing? Insights from a meta-
622 analysis. *Neuroimage*. **54**, 577–593 (2011).
- 623 57. R. V. Reichle, A. Tremblay, C. Coughlin, Working memory capacity in L2
624 processing. *Probus*. **28**, 29–55 (2016).
- 625 58. Council of Europe, *Common European Framework of Reference for Languages:*
626 *learning, teaching, assessment* (Cambridge University Press, 2001).
- 627 59. J. C. Raven, J. H. Court, *Raven's progressive matrices and vocabulary scales*
628 (Oxford: Oxford psychologists Press, 1998).

- 629 60. B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Van Der
630 Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, A.
631 M. Dale, Whole brain segmentation: Automated labeling of neuroanatomical
632 structures in the human brain. *Neuron*. **33**, 341–355 (2002).
- 633 61. M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K.
634 Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, S. M. Smith, D. C. Van Essen,
635 A multi-modal parcellation of human cerebral cortex. *Nature*. **536**, 171–178 (2016).
- 636 62. C. R. Buchanan, M. E. Bastin, S. J. Ritchie, D. C. Liewald, J. W. Madole, E. M.
637 Tucker-Drob, I. J. Deary, S. R. Cox, The effect of network thresholding and
638 weighting on structural brain networks in the UK Biobank. *Neuroimage*. **211**, 116443
639 (2020).
- 640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672

673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706

Supplementary Materials for

White matter plasticity during second language learning within and across hemispheres

Xuehu Wei, Thomas C. Gunter, Helyne Adamson, Matthias Schwendemann, Angela D.

Friederici, Tomás Goucha, and Alfred Anwander

xuehuwei@cbs.mpg.de

This file includes:

Supplementary Methods

Supplementary Figures S1 to S5

Supplementary Tables S1 to S2

Supplementary Methods

Language learning procedure. In our study, we recruited a large group of young, healthy, native Arabic speakers to participate in a six-month intensive German language course to reach an intermediate level of proficiency (B1, first level of independent language proficiency). The second language (L2) teaching and assessment structure followed the Common European Framework of Reference for Languages (1, 2). According to the CEFR, the first intermediate level of independent language use (B1) can be reached after approximately 600 hours of language learning, which corresponds to a six-month intensive course. The courses have been structured in cooperation with the Herder-Institute of the University of Leipzig, Germany, which is specialized in research and teaching of German as an L2. The proficiency levels of this standard framework comprise six levels. Levels A1 and A2 represent the elementary use of the language for beginners. Levels B1 and B2 represent intermediate language levels. At the first intermediate level (B1) of German, a learner can understand the main points when clear, standard language is used and familiar topics related to work, school, leisure time, etc. are the focus. The learner can make a short statement to explain his views and plans. C1 and C2 are the highest possible levels. In our study, the participants underwent two phases (0-3 months and 3-6 months) of daily intensive classroom training in German (L2). The course combined classroom teaching using standard textbooks, complex naturalistic speaking and reading, and clear instruction in grammar and vocabulary. The course took place at the Max Planck Institute in Leipzig, in small groups of 12-15 students, 45 minutes per lesson, 5 lessons per day, 5 days per week. Three different professional teachers taught the classes in each group to increase the language input of the learners and to reduce instructional variance between the groups. Daily

707 homework was assigned by the teachers and consisted mainly of consolidating and reviewing
708 the topics taught. In addition, several Arabic-speaking student assistants helped the learners
709 with everyday issues so that they could concentrate fully on the language courses.

710 **Language proficiency test.** After 3 months and 6 months of learning, participants took a 90-
711 minute standardized second language proficiency test that assessed language comprehension
712 and production performance through four subtests in listening, reading, writing, and speaking
713 in German. Language acquisition was assessed in the first phase between 0 and 3 months with
714 the standardized A1 test and in the second phase between 3 and 6 months with the B1 language
715 test of the German Goethe Institute, which tests oral and written, receptive and productive skills
716 (i.e. listening, reading, speaking, and writing). An additional L2 Vocabulary Size Test (VST)
717 was taken by a subgroup of 41 participants (35 males) after 6 months of learning (3). This
718 receptive VST was developed at the Institute for Test Research and Test Development, Leipzig
719 (ITT-Leipzig, <http://www.itt-leipzig.de>, (4)). The students were tested on their knowledge of a
720 sample of the 3000 most common words required at this level. It measures in five sections, how
721 many of a sample of words belonging to a given frequency range are known (1000 most
722 frequent, 2000 most frequent, etc.). This results in a maximum score of 30 points per section,
723 which were added together.

724 **Transformation of the A1 and B1 language scores to a common scale.**

725 To estimate the L2 proficiency longitudinally and correlate it with the brain structural plasticity
726 during learning, scores from each language test at each time point were scaled to a common
727 scale following the Cambridge English Scale ([https://www.cambridgeenglish.org/exams-and-
728 tests/cambridge-english-scale](https://www.cambridgeenglish.org/exams-and-tests/cambridge-english-scale)). In our study, the progress scale was always divided into steps
729 of 5. The detailed conversion relationship between the test score and the common scale is shown
730 in [Supplementary Table S1](#).

731
732 **MRI data acquisition.** Structural and high-resolution diffusion-weighted MR images were
733 acquired on a 3 Tesla Prisma MRI system (Siemens Healthineers, Erlangen, Germany) with a
734 32-channel head coil with the following scanning parameters: Isotropic voxels resolution of 1.3
735 mm, 60 diffusion directions ($b = 1000 \text{ s/mm}^2$) and 7 images without diffusion weighting ($b = 0$
736 s/mm^2), TE = 75 ms, TR = 6 s, GRAPPA = 2, CMRR-SMS=2, 3 repetitions to improve the
737 signal-to-noise ratio, and 2 b_0 acquisitions with opposite phase encoding. The diffusion
738 sequence was repeated 3 times to increase the. For the anatomical segmentation, we acquired
739 quantitative multiparametric structural images with 1 mm resolution (5). The images were

740 preprocessed using the publicly available hMRI toolbox (<http://hmri.info>) and the quantitative
741 magnetization transfer (MT) images were used for the segmentation and parcellation steps.

742

743 **Connectivity analysis:**

744 **Diffusion MRI preprocessing.** Preprocessing of diffusion data was performed using the
745 FMRIB Software Library (FSL, <http://www.fmrib.ox.ac.uk/fsl>). Diffusion images were
746 corrected for susceptibility and eddy current induced distortions as well as head motion with
747 the FSL tools “topup” and “eddy” using optimized parameters matched to image resolution.
748 Imaging noise in the high-quality diffusion MRI data was minimized by combining the three
749 repetitions. No additional denoising algorithms were applied to minimize image blurring. The
750 optimized imaging settings allowed Gibbs ringing artifacts to be minimized and an additional
751 correction was not required. No additional intensity or bias field correction was applied. Finally,
752 the brain volume was masked from the background and the standard DTI contrast maps were
753 computed. The processed datasets were checked individually to exclude artifacts from the
754 acquisition or preprocessing. Finally, the voxel-wise fiber distribution for probabilistic
755 tractography was computed with up to 3 fiber directions per voxel using the FSL command
756 “bedpostX” (6).

757 **Surface segmentation.** For each participant, the structural connectome of each hemisphere was
758 computed as follows (see also [Supplementary Figure S1](#)). First, the cortical and white matter
759 surface, as well as the 5 subsections of the corpus callosum of each participant were generated
760 from the MT images using FreeSurfer 5.3 (<http://surfer.nmr.mgh.harvard.edu>, (7)). The white
761 matter surface was shifted 1 mm into the white matter using the FreeSurfer command
762 “mris_expand” to define robust seed and target regions for probabilistic tractography.

763 **Parcellation of the seed regions.** The cortical surface was divided into 180 regions in each
764 hemisphere using the multi-modal parcellation developed as part of the Human Connectome
765 Project (HCP) (8). Therefore, the atlas annotations were transformed into separate labels using
766 “mri_annotation2label“, and then mapped to each participant using “mri_label2label”. Finally,
767 the labeled white matter surface (shifted 1mm inside the white matter) was mapped to the
768 individual anatomical voxel space using “mri_label2vol” to generate a voxel-based definition
769 of parcellation corresponding to the cortical areas. The labeled cortical and CC regions were
770 registered to the diffusion images (FA contrast) using a rigid body registration using “flirt” and
771 applied to the labeled regions using nearest-neighbor interpolation. Using the corpus callosum
772 sections as seed and target regions in probabilistic tractography reduces the problem of spurious

773 white matter connections resulting from the tracking through the bottleneck of the corpus
774 callosum. This results in a more robust estimation of the inter-hemispheric connectivity.

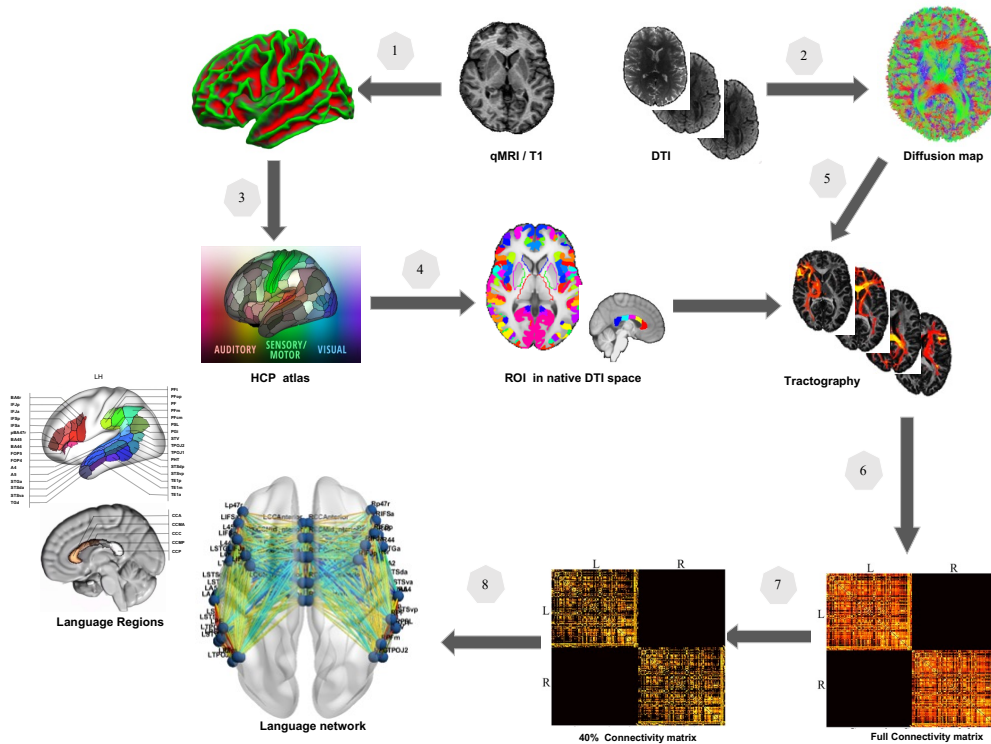
775 **Connectivity estimation.** All registered regions were used as seed areas for probabilistic
776 tractography (6) using “probtrackX” with default parameters. The structural connectivity
777 between all regions was computed representing the relative number of streamlines between all
778 pairs of regions. Those connectivity estimates are influenced by the local microstructural
779 properties of the pathway between the regions and integrate the local properties into one
780 connectivity estimate for a specific connection. The estimated connectivity values for all
781 regions (full HCP atlas) were logarithmically scaled and normalized by the size of the seed
782 region (log of the number of seeded streamlines) to build a connectivity matrix with normalized
783 values ranging from zero to one. Next, the values for both tracking directions were averaged
784 (from region A to region B, and from region B to region A). For each participant and each
785 hemisphere, we obtained the weighted symmetric connectome matrix ([Supplementary Figure](#)
786 [S1](#)).

787 **Network thresholding.** Additionally, we removed weak and noisy connections below a
788 predefined threshold (in the average matrix across all participants) as they cannot be estimated
789 reliably with tractography due to the limited sampling of the distribution that may result in
790 false-positive connections (9). This allowed the exclusion of connections that did not align with
791 the major fiber pathways in the human brain (10), and removed, e.g., connections between the
792 left parietal lobe and frontal CC regions that do not exist anatomically. To determine this
793 threshold, we increased the threshold in increments of 10% to create seven networks with
794 different densities. Network thresholding methods were shown to be able to disentangle
795 spurious and genuine connections (9). These networks ranged from a dense network that
796 contained 80% of all connections to a sparse network that included only the strongest 20% per
797 hemisphere. A threshold of 30% was found to reliably remove implausible false-positive
798 connections and still retain the major pathways for the network-based analysis. With this global
799 density threshold, 67% of the connections within the cortical language network (out of 528 per
800 hemisphere, $33 \times 32 / 2$) and 55% of all possible connections including the CC areas were retained.
801 The same network mask was applied to every individual participant. Finally, the 33 language
802 ROIs in each hemisphere were selected and the matrix was reduced to those elements for further
803 analysis.

804 **Network-based statistics.** We used network-based statistics (NBS) (11) to identify
805 subnetworks with systematic structural changes. NBS is a method to control for the family-wise

806 error rate when testing each connection in the network by using the extent to which the edges
807 are connected. Therefore, all connected components that were present in the set of supra-
808 threshold connections (T-threshold = 3.3) were identified and the number of connections was
809 stored. To estimate the significance of each component, NBS performed a nonparametric
810 permutation test (K = 5000 permutations). At each permutation, the group to which each
811 participant belonged was randomly exchanged, the same threshold is applied to create the set
812 of connections above the threshold for each K permutation, and then the statistical test was
813 recalculated and the size of the largest component m in the set of supra-threshold connections
814 was stored. The p-value of each connected component of size m was then estimated by
815 searching for the proportion of permutations for which the maximal component size was greater
816 than m , and was then normalized by K. In this way, the NBS attempts to utilize the presence of
817 any structure exhibited by the connections comprising the effect or contrast of interest to yield
818 greater power than what is possible by independently correcting the p-values computed for each
819 link using a generic procedure to control the FWE (11, 12).

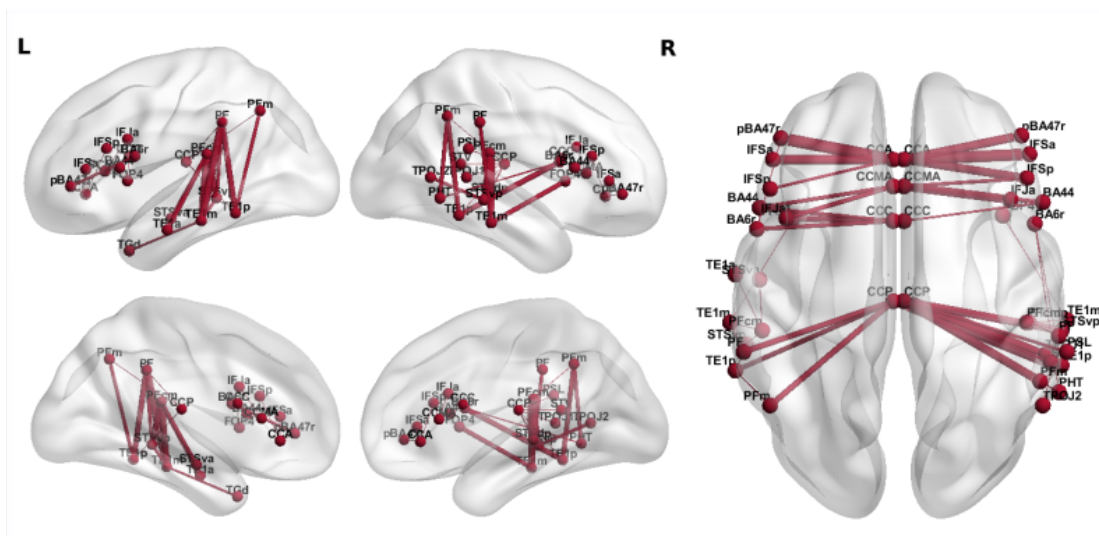
820 **Supplementary Figures:**



821
822

823 **Figure S1: Processing pipeline of structural connectome construction.** 1. White
824 white matter surface was generated by a segmentation of the anatomical MRI. 2.
825 Preprocessing of the diffusion MRI. 3. Atlas-parcellation of cortical regions in the
826 native brain surface. 4. Registration of the parcellated surface to native diffusion
827 space. 5. Tractography using seed regions in diffusion space. 6. Whole brain
828 network computation. 7. Network thresholding. 8. Extraction of the language
829 network.

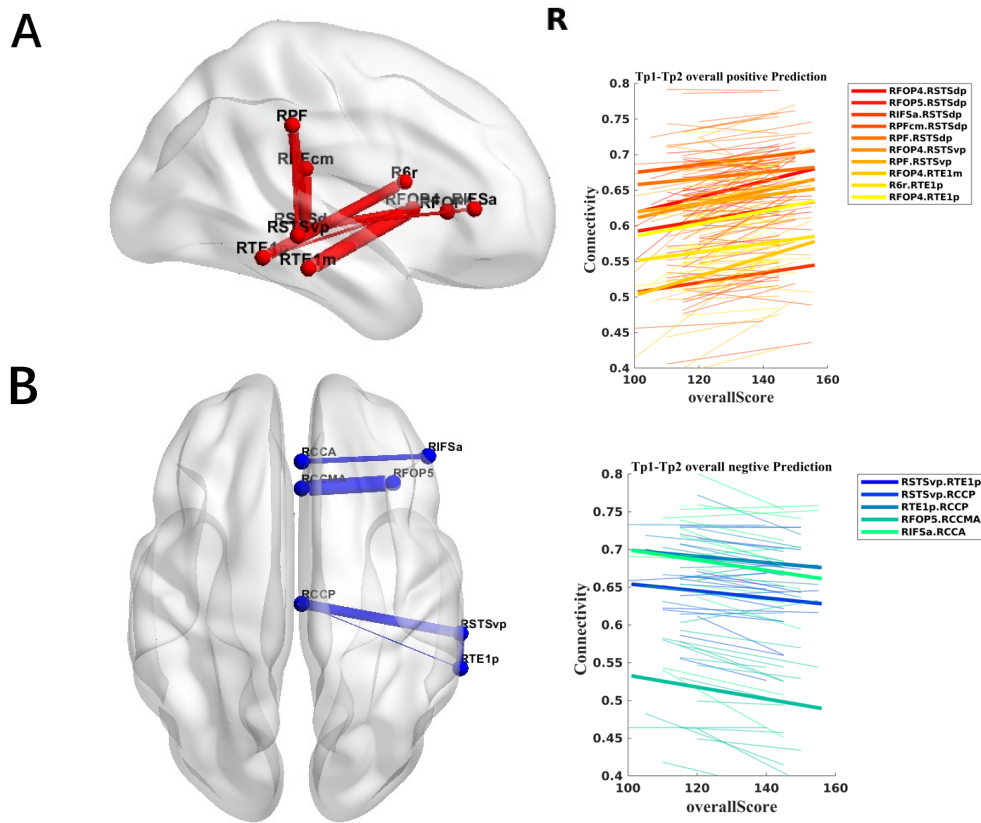
830



831

832
833

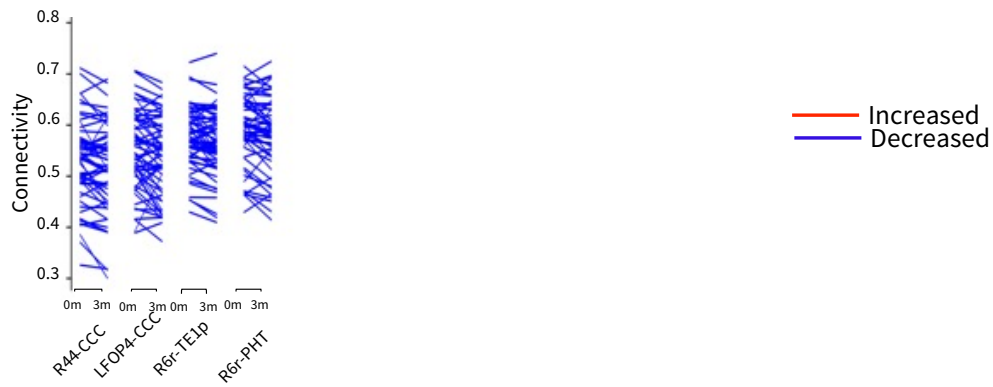
Figure S2: Longitudinal network changes across 3 learning time points.
($p < 0.05$ NBR corrected).



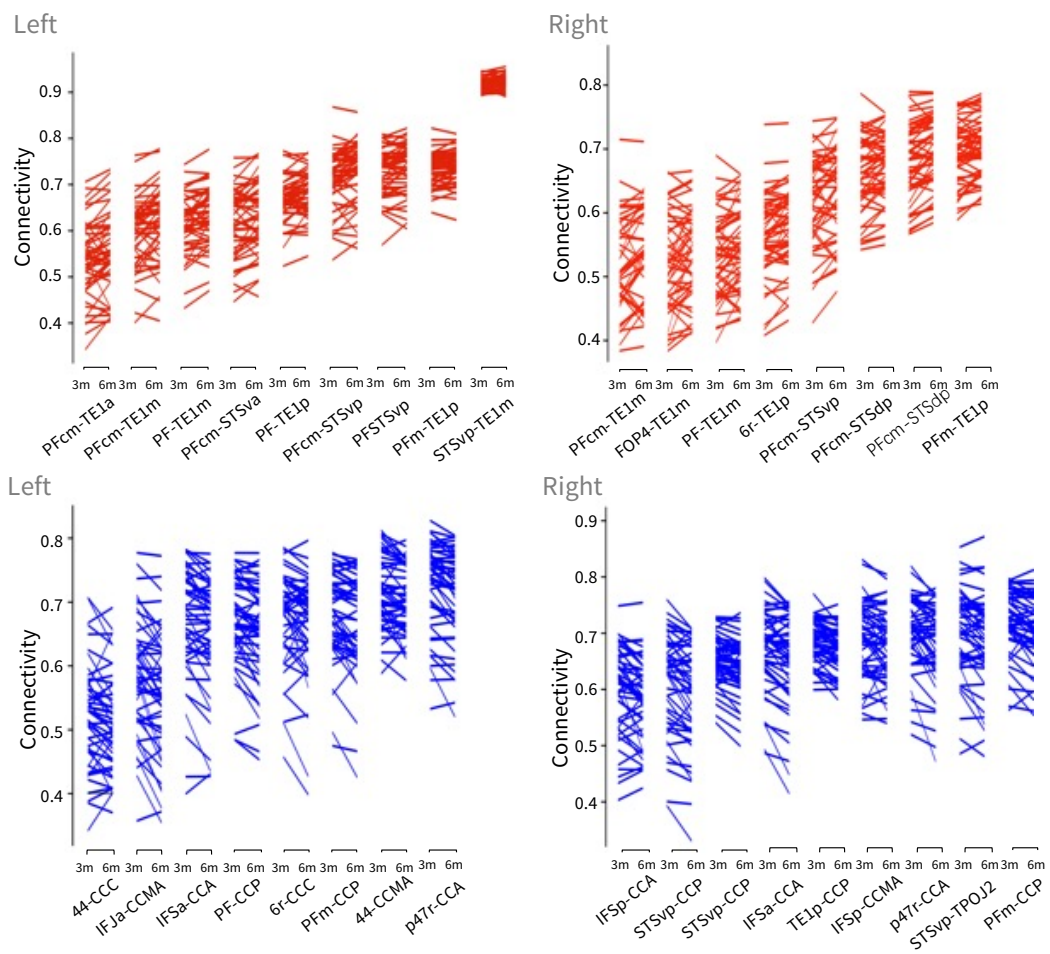
834
835
836
837
838
839
840

Figure S3: Subgroup with high productive vocabulary scores in the B1 test.
The correlation between the changes in connectivity and L2 proficiency from three to six months of L2 learning. Positive (A) and negative (B) correlation between L2 performance and changes in subnetworks. Note that the connection between the right temporal lobe and the frontal lobe follows the right arcuate fascicle as shown in Figure 3 of the main manuscript.

Baseline to 3months



3months to 6 months



841

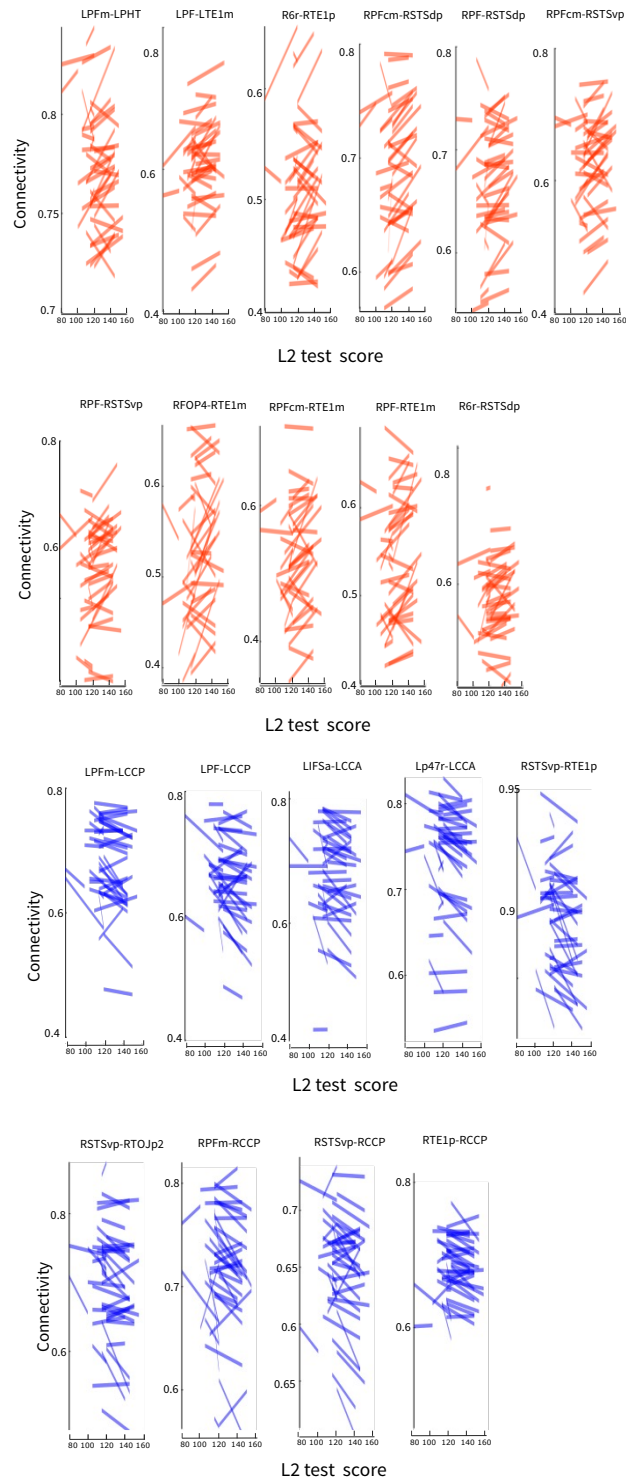
842

843

844

845

Figure S4: Individual data. Subnetworks with longitudinal increased and decreased connectivity in different L2 learning periods of each participant and each connection within the networks which showed significant changes.



846

847

848

849

850

851

852

Figure S5: Individual data. Longitudinal changes of the connectivity values in relation to the progress in the language test from 3 months to 6 months of learning (normalized scores).

853 **Supplementary Table 1:**

854

Common Scale	A1	B1
40	0-19.9	
60	20-39.9	
80	40-59.9	0-19.9
100	60-67.4	20-39.9
105	67.5-74.9	-
110	75-82.4	-
115	82.5-89.9	-
120	90-100	40-59.9
140		60-67.4
145		67.5-74.9
150		75-82.4
155		82.5-89.9
160		90-100

855

856 **Supplementary Table 1:** Transformation of the A1 and B1 language test scores to
 857 a common scale. The standardized tests require 60% to successfully pass the exam
 858 and are less quantitative below this threshold. Therefore, the conversion follows
 859 coarser discretization steps of 10 below 60%.

860

861 **Supplementary Table 2:**

HCP Atlas	ROI Name	Area Description	Region
4	BA44	Broca Area 44	IFG
75	BA45	Broca Area 45	IFG
78	BA6r	rostral Broca Area 6	IFG
79	IFJa	anterior Inferior Frontal Junction	IFG
80	IFJp	posterior Inferior Frontal Junction	IFG
81	IFSp	posterior Inferior Frontal Sulcus	IFG
82	IFSa	anterior Inferior Frontal Sulcus	IFG
108	FOP4	Frontal OPercular area 4	IFG
169	FOP5	Frontal OPercular area 5	IFG
171	pBA47r	posterior Broca Area 47 rostral	IFG
76	47l	Broca Area 47 lateral	IFG
105	PFcm	Area PFcm	IPL
116	PFt	Area PFt	IPL
147	PFop	Area PF opercular	IPL
148	PF	Area PF Complex	IPL
149	PFm	Area PFm Complex	IPL
150	PGi	Area PGi	IPL

25	PSL	PeriSylvian Language area	IPL
140	TPOJ2	Temporo Parieto Occipital Junction 2	IPL
123	STGa	Superior Temporal Gyrus anterior	TL
125	A5	Auditory 5 Complex	TL
128	STSda	Superior Temporal Sulcus dorsal anterior	TL
129	STSdp	Superior Temporal Sulcus dorsal posterior	TL
137	PHT	Area PHT	TL
175	A4	Auditory 4 Complex	TL
176	STSva	Superior Temporal Sulcus ventral anterior	TL
130	STSvp	Superior Temporal Sulcus ventral posterior	TL
131	TGd	Temporal pole dorsal	TL
132	TE1a	Temporal area 1 anterior	TL
177	TE1m	Temporal area 1 middle	TL
133	TE1p	Temporal area 1 posterior	TL
139	TPOJ1	Temporo Parieto Occipital Junction 1	TL
28	STV	Superior Temporal Visual area	TL
--	CCa	Corpus Callosum anterior	aCC
--	CCma	Corpus Callosum middle anterior	aCC
--	CCc	Corpus Callosum central	aCC
--	CCmp	Corpus Callosum middle posterior	pCC
--	CCp	Corpus Callosum posterior	pCC

862
863
864
865
866
867

868

869

870

871

872

873

874

875

Supplementary Table 2: Labels for each language region. Inferior Frontal Gyrus: IFG, Temporal Lobe: TL, Inferior Parietal Lobe: IPL, anterior/posterior Corpus Callosum: aCC / pCC.

876 **SI References**

- 877 1. Council of Europe, *Common European Framework of Reference for Languages:*
878 *learning, teaching, assessment* (Cambridge University Press, 2001).
- 879 2. D. Little, The Common European Framework of Reference for Languages: Content,
880 purpose, origin, reception and impact. *Language Teaching*. **39**, 167–190 (2006).
- 881 3. E. Tschirner, *Aligning frameworks of reference in language testing: The ACTFL*
882 *proficiency guidelines and the common European framework of reference for*
883 *languages*. (Stauffenburg-Verlag, 2012).
- 884 4. E. Tschirner, Examining the validity and reliability of the ITT vocabulary size tests.
885 Research Papers in Assessment. *Univ. Leipzig*. **3** (2021).
- 886 5. N. Weiskopf, J. Suckling, G. Williams, M. M. Correia M., B. Inkster, R. Tait, C. Ooi,
887 E. T. Bullmore T., A. Lutti, Quantitative multi-parameter mapping of R1, PD*, MT,
888 and R2* at 3T: A multi-center validation. *Front. Neurosci.* **7**, 95 (2013).
- 889 6. T. E. J. Behrens, H. J. Berg, S. Jbabdi, M. F. S. Rushworth, M. W. Woolrich,
890 Probabilistic diffusion tractography with multiple fibre orientations: What can we
891 gain? *Neuroimage*. **34**, 144–155 (2007).
- 892 7. A. M. Dale, B. Fischl, M. I. Sereno, Cortical surface-based analysis: I. Segmentation
893 and surface reconstruction. *Neuroimage*. **9**, 179-194 (1999).
- 894 8. M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K.
895 Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, S. M. Smith, D. C. Van Essen,
896 A multi-modal parcellation of human cerebral cortex. *Nature*. **536**, 171–178 (2016).
- 897 9. C. R. Buchanan, M. E. Bastin, S. J. Ritchie, D. C. Liewald, J. W. Madole, E. M.
898 Tucker-Drob, I. J. Deary, S. R. Cox, The effect of network thresholding and
899 weighting on structural brain networks in the UK Biobank. *Neuroimage*. **211**, 116443
900 (2020).
- 901 10. K. H. Maier-Hein, *et al.*, The challenge of mapping the human connectome based on
902 diffusion tractography. *Nat. Commun.* **8**, 1349 (2017).
- 903 11. A. Zalesky, A. Fornito, E. T. Bullmore, Network-based statistic: Identifying
904 differences in brain networks. *Neuroimage*. **53**, 1197–1207 (2010).
- 905 12. L. García-Pentón, A. Pérez Fernández, Y. Iturria-Medina, M. Gillon-Dowens, M.
906 Carreiras, Anatomical connectivity changes in the bilingual brain. *Neuroimage*. **84**,
907 495–504 (2014).
- 908