JNeurosci
THE JOURNAL OF NEUROSCIENCE

JNeurosci Accepted Manuscript

Research Articles: Behavioral/Cognitive

Converging evidence for differential specialisation and plasticity of language systems

This Early Release article has been peer-reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.jneurosci.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

Converging evidence for differential specialisation and plasticity of language systems

Abbreviated title: Specialisation and plasticity of language systems

Kshipra Gurunandan*[1], Jaione Arnaez-Telleria[1], Manuel Carreiras[1,2,3] and Pedro M. Paz-Alonso*[1]

[1] BCBL Basque Center on Cognition, Brain and Language, 20009 Donostia-San Sebastian, Spain
[2] IKERBASQUE Basque Foundation for Science, 48013 Bilbao, Spain
[3] Department of Basque Language and Communication, EHU/UPV, 48015 Bilbao, Spain

*Correspondence should be addressed to: Kshipra Gurunandan, k.gurunandan@bcbl.eu and Pedro M. Paz-Alonso, p.pazalonso@bcbl.eu

K.G., J.A.-T., M.C. and P.M.P-A. designed research; K.G. and J.A.-T. performed research; K.G. and P.M.P-A. analysed data and wrote the paper; M.C. provided comments on the manuscript.

The manuscript includes **5** figures and **1** table.
Number of words: Abstract (221 words), Introduction (676 words), Discussion (**1961** words).

Abstract

Functional specialisation and plasticity are fundamental organising principles of the brain. Since the mid-1800s, certain cognitive functions have been known to be lateralised, but the provenance and flexibility of hemispheric specialisation remain open questions. Language is a uniquely human phenomenon that requires a delicate balance between neural specialisation and plasticity, and language learning offers the perfect window to study these principles in the human brain. In the current study, we conducted two separate functional MRI experiments with male and female language learners, one cross-sectional and one longitudinal, involving distinct populations and languages, and examined hemispheric lateralisation and learning-dependent plasticity of three language systems: reading, speech comprehension and verbal production. A multi-pronged analytic approach revealed a highly consistent pattern of results across the two experiments, showing that (i) in both native and non-native languages, while language production was left-lateralised, lateralisation for language comprehension was highly variable across individuals, and (ii) with increasing non-native language proficiency, reading and speech comprehension displayed substantial changes in hemispheric dominance, with languages tending to lateralise to opposite hemispheres; while production showed negligible change and remained left-lateralised. These convergent results shed light on the long-standing debate of neural organisation of language by establishing robust principles of lateralisation and plasticity of the main language systems. Findings further suggest involvement of the sensorimotor systems in language lateralisation and its plasticity.

Significance Statement

The human brain exhibits a remarkable ability to support a vast variety of languages that may be acquired at different points in the lifespan. Language is a complex construct involving linguistic as well as visual, auditory, and motor processes. Using functional MRI, we examined hemispheric specialisation and learning-dependent plasticity of three language systems — reading, speech comprehension and verbal production — in cross-sectional and longitudinal experiments in language learners. A multi-pronged analytic approach revealed converging evidence for striking differences in hemispheric specialisation and plasticity between the language systems. The results have major theoretical and practical implications for our understanding of fundamental principles of neural organisation of language, language testing and recovery in patients, and language learning in healthy populations.

Introduction

Functional specialisation in the brain is a well-established principle of neural organisation, but studies of atypical development suggest dramatic potential for neural plasticity (Payne and Lomber 2001; Bavelier and Neville 2002). While the capacity for neural reorganisation decreases with age, it does not disappear completely, and adult neural plasticity is essential for learning and maintaining new information or behaviours (Kleim and Jones 2008). The human propensity for language requires a delicate balance between neural specialisation and capacity for re-organisation, making language learning the ideal candidate for examination of specialisation and plasticity in the human brain.

Language typically activates a fronto-temporo-parietal network (Skeide and Friederici, 2016; Hagoort, 2019), and has long been thought to be predominantly left-lateralised (Broca, 1863; Dax, 1863). However, the right hemisphere appears to be capable of taking over or supporting language function if needed, as seen in cases of language recovery after left-hemisphere damage (Papanicolaou et al., 1987; Boatman et al., 1999; Duffau et al., 2002, 2003; Hope et al., 2017) and in language learning (Vingerhoets et al., 2003; Park et al., 2012). It is thus unclear whether the left hemisphere is indeed specialised for language as is broadly accepted, with the right hemisphere playing at best a supporting role (Vigneau et al., 2010), or whether hemispheric dominance is more variable across individuals, as suggested by the larger than expected prevalence of language deficits following right hemisphere brain surgery (Vilasboas et al., 2017).

Language is a complex construct involving multi-level representations that can be processed visually (reading), auditorily (listening) or by motor production (speaking/writing), and cumulative evidence points to these functions lateralising differently. Auditory language has been found to be bilateral in infants (Dehaene-Lambertz et al., 2002; Perani et al., 2011), with either no increase in lateralisation from childhood to adulthood (Lidzba et al., 2011), increasing left-lateralisation (Ahmad et al., 2003), or increasing right-hemisphere involvement (Booth et al.,

2000), and a meta-analysis of auditory comprehension studies suggested that any left-lateralisation from childhood to adulthood increases more slightly and gradually than previously thought (Enge et al., 2020). On the other hand, there is little evidence to suggest that language production is anything but left-lateralised (Gaillard et al., 2003; Szaflarski et al., 2006; Lidzba et al., 2011).

Language learning is known to change the pattern of neural activation for language. Studies comparing bilinguals and monolinguals consistently find differences in activation between them, with bilinguals typically exhibiting greater right hemispheric involvement in comprehension tasks (e.g. Kovelman et al., 2008; Horowitz-Kraus et al., 2015). However, it is uncertain whether this increased right hemispheric involvement merely modulates the magnitude of left-lateralisation, or whether it is significant enough to constitute a change in hemispheric dominance. Further, are differences in lateralisation between monolinguals and bilinguals due to developmental differences or is hemispheric dominance in fact plastic even into adulthood? Few neuroimaging studies have looked into ecologically-valid adult language learning, but findings indicate that language learning in adults involves structural changes in cortical thickness and connectivity that could indeed support shifts in lateralisation (Mårtensson et al., 2012; Schlegel et al., 2012; Xiang et al., 2015), suggesting that lateralisation, at least for comprehension, may be susceptible to learning-dependent changes.

We conducted two fMRI experiments, one cross-sectional and one longitudinal, with immersed late language learners, and examined lateralisation of reading, speech comprehension, and verbal production in their native (L1) and non-native (Ln) languages, and how this changed with increasing Ln proficiency. To test both replicability and generalisability of findings, the two experiments were contrasted on several factors such as early language experience of the participants (monolingual vs bilingual) and the language currently being learnt, and the L1-Ln pairs in the two experiments had contrasting degrees of overlap in language families, phonology, and orthography. We hypothesised that (i) lateralisation of comprehension would be more

83 variable across individuals but production would be left-lateralised, and (ii) with increasing

84 language proficiency, comprehension may display changes in hemispheric dominance, while

85 production would remain left-lateralised. We further expected that L1-Ln associations would

86 change with increasing Ln proficiency, and that the pattern of changes would differ across the

87 language systems.

88

Materials and Methods

89

Participants

90

Experiment I: basic vs advanced level language learners (cross-sectional)

91

92 The final experiment sample consisted of 29 right-handed native Spanish adults (mean

93 age=43.7 ± 9.7 years; 15 female) studying Basque in the same language school at either the basic

94 (A2 level, n=14) or advanced level (C1 level, n=15). The proficiency levels correspond to those

95 specified by the Common European Framework of Reference for Languages (CEFR). Participants

96 were from the Basque Country, Spain; they grew up primarily exposed to Spanish at home and in

97 school, with little early Basque exposure, and had limited knowledge of English or other

98 languages (no difference between groups, p=0.83). The two groups of learners were matched on

99 age, gender, IQ and Spanish proficiency (Table 1). Data from 5 other participants was discarded

100 due to excessive head motion during MRI scanning and these were not counted in the final

101 sample.

102

Experiment II: intermediate language learners (longitudinal)

103

104 The final experimental group consisted of 19 right-handed native Spanish adolescents

105 (mean age=17.2 ± 0.6 years; 16 female) taking part in a 3-month English immersion-style after-

106 school programme for B1 level students. Participants were from the Basque Country, Spain; they

107 were native speakers of Spanish and acquired Basque in school (AoA=2.6 ± 2.06 years). The

108   medium of instruction in school was Spanish/Basque; English was learnt as a foreign language,

109   with little exposure outside of classes. The students had intermediate English proficiency (Table

110   1). Data from 5 other participants was discarded due to excessive head motion during MRI

111   scanning and these were not counted in the final sample.

112

113   **Experiments I & II**

114   In both experiments, language proficiency was assessed using picture-naming tasks — an

115   adaptation of the Boston Naming Test (Kaplan et al., 1983) controlled for cognates across Spanish,

116   Basque and English. Participant groups in Experiment I differed significantly in their Basque

117   proficiency, and participants in Experiment II exhibited significant increase in English

118   proficiency after language training (Table 1). All participants had normal or corrected-to-normal

119   vision, and no history of neurological or psychiatric disorders. In compliance with the ethical

120   regulations established by the BCBL Ethics Committee and the guidelines of the Helsinki

121   Declaration, all participants gave written informed consent prior to taking part in the

122   experiment, and received monetary compensation for their participation.

123

124   fMRI Task

125   Inside the MRI scanner, participants performed two tasks: a comprehension and a

126   production task. The order of tasks was counter-balanced across participants.

127

128   **Language Comprehension Task**

129   The participants performed semantic animacy judgement (living/non-living) with single-

130   word text and speech stimuli in each of their languages. Participants were instructed to fixate on

131   a white cross in the middle of a black screen, and on presentation of stimuli, to indicate their

132   responses as quickly and as accurately as possible via button presses (counter-balanced across

participants) using their dominant (right) hand. Stimuli were high frequency, concrete, imageable nouns (e.g. house, dog, table) with an even split between living and non-living items. Visual stimuli were presented in white letters on a black screen and were 5-8 letters long. Auditory stimuli were presented through headphones and lasted an average of 565 ms (s.d.=86 ms). Each run had 48 stimuli with inter-mixed reading and listening trials. The fMRI design was event-related with six/four runs (Experiment I: 2 languages x 3 runs; Experiment II: 2 languages x 2 runs). To avoid language-switching, the languages were separated and their order was counterbalanced across participants.

Language Production Task

The participants performed a paced form of the semantic verbal fluency task in each language. Participants were instructed to fixate on a white cross in the middle of a black screen and respond overtly to semantic category words (e.g. fruits, animals, clothes) presented on the screen. Each word was displayed eight times, each requiring a novel response, or failing this, an overt response saying "pass" in the relevant language. Fluency was scored as the percentage of valid answers out of eight possible responses for each category. Repetitions, inflections of the same word and erroneous responses were removed, and responses were scored only for correctness and not accent or pronunciation. In the control condition, participants repeated the word presented on the screen. The task had a block design with two runs per language, each run containing eight semantic categories. To avoid language-switching, the languages were separated and their order was counterbalanced across participants.

MRI Data Acquisition

Whole-brain MRI data was acquired using a 3-T Siemens Magnetom Trio whole-body MRI scanner and a 32-channel head coil at the Basque Center on Cognition, Brain and Language

158   (BCBL). Padded headphones were used to dampen background scanner noise and enable clear

159   transmission of the auditory stimuli. Participants viewed the print stimuli on a screen via a

160   mirror mounted on the head coil. To limit head movement, the head coil was padded with foam

161   and participants were asked to remain as still as possible. Structural T1-weighted images were

162   acquired with a MPRAGE sequence with TR=2530 ms, TE=2.97 ms, inversion time=1100 ms, FA=7°,

163   FoV=256 x 256 mm, 176 slices and voxel size=1 mm^3.

164

165   **Language Comprehension Task**

166   Functional MRI was acquired in the course of six/four separate runs using a gradient-echo

167   echo-planar pulse sequence with the following parameters: TR 2000 ms, TE 30 ms, 32 axial slices

168   with a 3.4 x 3.4 x 4 mm voxel resolution, 0% inter-slice gap, flip angle (FA)=80°, field of view

169   (FoV)=220 mm, 64 x 64 matrix. 186 volumes were collected for each of the functional runs. Prior

170   to each scan, four volumes were discarded to allow for T1-equilibration effects. To improve

171   estimation of the resting baseline in functional analyses, functional runs contained three silent

172   fixation periods of 20 s each. Within each functional run, the order of the trials (reading and

173   listening conditions) and the inter-trial intervals of variable duration corresponding to the

174   baseline MR frames (30% of total collected functional volumes) were determined by an algorithm

175   designed to maximise the efficiency of the recovery of the blood oxygen level dependent

176   response (optseq2, Dale 1999).

177

178   **Language Production Task**

179   Functional MRI was acquired in the course of four separate runs using a gradient-echo

180   echo-planar pulse sequence with the following parameters: TR 3000 ms, TE 25 ms, 43 axial slices

181   with a 3.0 x 3.0 x 3.0 mm voxel resolution, 10% inter-slice gap, flip angle (FA)=90°, field of view

182   (FoV)=192 mm, 64 x 64 matrix. 240 volumes were collected for each of the functional runs. Prior

183   to each scan, four volumes were discarded to allow for T1-equilibration effects.

184

185   ## MRI Data Analysis

186   ### Preprocessing

187   Standard SPM8 (Penny et al., 2011) preprocessing routines and analysis methods were

188   employed. Images were first corrected for differences in timing of slice acquisition and then

189   realigned to the first volume using rigid-body registration. Each subject's functional volumes

190   were spatially smoothed with a 4-mm full-width half-maximum (FWHM) Gaussian kernel. Next,

191   motion parameters obtained from realignment were used to inform a volume repair procedure

192   (ArtRepair, Mazaika et al., 2009) that identified bad volumes on the basis of scan-to-scan

193   movement (>1 mm) and signal fluctuations in global intensity (>1.3%), and corrected bad volumes

194   via interpolation between the nearest non-repaired scans. Data from subjects requiring more

195   than 20% of volumes to be repaired were discarded. The number of corrected volumes was similar

196   between groups (Experiment I: comprehension task $p=0.34$, production task $p=0.63$) and scans

197   (Experiment II: comprehension task $p=0.75$, production task $p=0.46$). After volume repair,

198   functional volumes were co-registered to the T1 images using 12-parameter affine

199   transformation and spatially normalised to the MNI space by applying non-linear transforms

200   estimated by deforming the MNI template to each individual's structural volume (Ashburner and

201   Friston, 2005). During normalisation, the volumes were sampled to 3-mm cubic voxels. The

202   resulting volumes were then spatially smoothed with a 7-mm FWHM Gaussian kernel. Finally,

203   time series were temporally filtered to eliminate contamination from slow frequency drift (high-

204   pass filter with cut-off period of 128 s).

205

Subject level analyses

Statistical analyses were performed on individual subject data using the general linear model (GLM). fMRI time series data were modelled by a series of impulses convolved with a canonical haemodynamic response function. Six motion parameters for translation (x, y, z) and rotation (yaw, pitch, roll) were included as covariates of non-interest in the GLM. In the event-related-design comprehension task, each trial was modelled as an event, time-locked to the onset of the presentation of each stimulus, and error responses were modelled separately. In the block-design production task, each trial was modelled as an epoch of 31 s each, time-locked to the beginning of the presentation of each block. The remaining functions were used as covariates in the GLM, along with a basic set of cosine functions that high-pass filtered the data, and a covariate for session effects. The least-squares parameter estimates of the height of the best-fitting canonical HRF for each experimental condition were used in pairwise contrasts.

Laterality analyses

For every subject, lateralisation of activation in the classical language network regions was calculated for each task x language. Laterality is typically quantified as a normalised ratio of left and right hemisphere contributions, ranging between +1 (fully left-lateralised activation) and -1 (fully right-lateralised activation). Each subject's whole-brain t-maps were masked with anatomical language regions from the AAL atlas (Tzourio-Mazoyer et al., 2002) — six bilateral regions from classical language models (Friederici 2012; Hagoort 2013): inferior frontal gyrus (IFG) *pars orbitalis*, IFG *pars triangularis*, IFG *pars opercularis*, superior temporal gyrus (STG), middle temporal gyrus (MTG), and inferior parietal lobule (IPL). Since laterality indices are highly threshold-dependent, in line with the latest recommendations (Bradshaw et al., 2017), a threshold-independent bootstrapping method was used to calculate the laterality index using the LI-toolbox (Wilke and Lidzba, 2007), in which 10,000 indices were iteratively calculated at

different thresholds, yielding a robust mean laterality index. Three analyses were carried out to examine proficiency-dependent plasticity of (i) L1-Ln correlation, (ii) hemispheric dominance, and (iii) modality clustering. L1-Ln correlations were calculated for each group x task, and Cohen's *q* was used to quantify the difference in L1-Ln correlation between basic/advanced proficiency and before/after training in each modality. To examine hemispheric dominance, a lateralised dissociation index was calculated such that:

$$Lateralised\ Dissociation\ Index\ = |\ LI_{L1}\ -\ LI_{Ln}|\ *\ hem$$

$$hem = \begin{cases} 1 \text{ if opposite lateralisation} \\ -1 \text{ if same lateralisation} \end{cases}$$

i.e. the absolute difference between laterality indices for each language and a factor *hem* to code whether the two languages were lateralised to the same or opposite hemispheres. Positive values indicated that languages were lateralised to opposite hemispheres, while negative values indicated that the languages were lateralised to the same hemisphere. Cohen's *d* was used to measure the magnitude of proficiency-dependent change in each modality: difference between medians in cross-sectional Experiment I, and difference in repeated measures in longitudinal Experiment II. To examine the modality-wise clustering of the joint L1-Ln distribution, 85% data ellipses were plotted for each modality, and the joint distribution difference (JDD) between any two modalities was calculated as:

$$Joint\ Distribution\ Difference = \frac{distance\ between\ centroids}{maximum\ distance} * \frac{angle\ between\ major\ axes}{maximum\ angle}$$

i.e. standardised difference between the bivariate L1-Ln group means and difference between joint spread of the data. This index lies between 0 and 1, with higher values indicating greater

difference between modalities. The maximum Euclidean distance between centroids was considered to be 1 for laterality data, and maximum angle between the axes is 90°. Proficiency-group differences were measured in terms of percentage difference in the difference index.

Results

In-Scanner Behavioural Performance

Experiment I: basic vs advanced level language learners (cross-sectional)

A series of mixed-model analyses of variance (ANOVAs) were conducted on the behavioural measures of the fMRI tasks: comprehension task accuracy, production task fluency, and comprehension task reaction times (Figure 1A). The comprehension task accuracy ANOVA with between-subjects factor Group (basic, advanced) and within-subject factors Language (L1, Ln) and Modality (reading, speech) showed a significant Group x Language interaction ($F_{(1,26)}=16.18$, $p=0.0004$). The production task fluency ANOVA with between-subjects factor Group (basic, advanced) and within-subject factor Language (L1, Ln) also showed a significant Group x Language interaction ($F_{(1,23)}=31.36$, $p=0.00001$). Post-hoc simple-effect analyses (two-sample t-tests) of these Group x Language interactions showed that the advanced proficiency group had significantly higher Ln task accuracy than the basic proficiency group in both comprehension ($t(18.08)=3.20$, $p=0.002$, one-sided) and production ($t(22.28)=5.502$, $p=0.000008$, one-sided), but there was no significant difference between groups in L1 task accuracy (comprehension: $t(25.48)=-0.93$, $p=0.360$, two-sided, production: $t(21.683)=1.03$, $p=0.31$, two-sided). Finally, the ANOVA for comprehension task reaction times showed a main effect of Language, and both groups were significantly slower in their Ln than their L1 ($F_{(1,26)}=40.41$, $p=0.000001$).

Experiment II: intermediate language learners (longitudinal)

A series of repeated-measures ANOVAs were conducted on the behavioural measures of the fMRI tasks: comprehension task accuracy, production task fluency, and comprehension task reaction times (Figure 1B). The comprehension task ANOVAs with 3 within-subject factors Training (before, after), Language (L1, Ln), and Modality (reading, speech), showed main effects of Language (L1 > Ln, $F(1,17)=338.64$, $p=0.000000000001$) and Modality (reading > speech, $F(1,17)=30.05$, $p=0.00004$) on task accuracy. The production task fluency ANOVA with 2 within-subject factors Training (before, after), and Language (L1, Ln) showed a main effect of Language (L1 > Ln, $F(1,15)=146.01$, $p=0.000000004$). The comprehension task reaction times ANOVA revealed a significant Training x Language interaction ($F(1,17)=5.48$, $p=0.031$). Post-hoc simple-effect analyses (paired t-tests) showed that reaction times decreased significantly after training in Ln ($t(17)=2.83$, $p=0.006$, one-sided), but not in L1 ($t(17)=0.21$, $p=0.836$, two-sided).

Language lateralisation

1. Lateralisation in comprehension and production

Laterality indices were calculated for the language network regions in each task and language using a threshold-free method, with values between +1 (left lateralisation) and -1 (right lateralisation). In both experiments, Wilcoxon signed-rank tests of paired samples revealed significant differences between each of the modalities. Comprehension and production displayed robust differences in lateralisation, with significant differences between both reading and verbal production (Experiment I: $W=421$, $p=0.000000000006$, Experiment II: $W=1099$, $p=0.0000000004$) as well as between speech comprehension and verbal production (Experiment I: $W=824$, $p=0.000002$, Experiment II: $W=729.5$, $p=0.000000000000006$). Reading and speech comprehension also differed significantly (Experiment I: $W=1998.5$, $p=0.021$, Experiment II: $W=3840$, $p=0.012$). In reading and speech comprehension, lateralisation was highly variable and indices spanned the full range of possible values between the two languages, while verbal production was clearly left-lateralised.

At the group level, comprehension appeared bilateral and production was left-lateralised. This result was consistent across the cross-sectional and longitudinal experiments (Figure 2).

2. Learning-dependent changes in lateralisation

To examine patterns of learning-dependent changes in lateralisation while accounting for the high inter-individual variability across tasks and languages, L1 lateralisation was used as a baseline for each subject's Ln lateralisation, and the linear association between L1 and Ln was assessed using Pearson's *r*. In lower proficiency learners, L1 and Ln lateralised similarly, regardless of left/right lateralisation. However, with increasing proficiency, this pattern reversed for comprehension, and L1 and Ln lateralised to opposite hemispheres. This learning-dependent change was not observed in verbal production (Figure 3). Cohen's *q* was used to quantify the proficiency-dependent change in L1-Ln correlation for each task, confirming that, across both studies, learning-dependent change in lateralisation was large in reading comprehension, medium in speech comprehension, and small in verbal production.

To examine whether increasing proficiency involved changes in hemispheric dominance for each modality, lateralised dissociation indices were calculated for each subject such that absolute values indicated the magnitude of L1-Ln difference, and direction (i.e. positive or negative) indicated whether the languages were lateralised to same or opposite hemispheres (positive=opposite hemispheres, negative=same hemisphere). There was a significant proficiency-related increase in absolute dissociation between L1 and Ln lateralisation across modalities (Experiment I: Mann-Whitney U tests: across modalities: $W = 584.5$, $p = 0.013$; reading comprehension: $W = 41$, $p = 0.007$; speech comprehension: $W = 88$, $p = 0.579$; verbal production: $W = 53.5$, $p = 0.022$; Experiment II: Wilcoxon signed-rank tests: across modalities: $V = 374.5$, $p = 0.023$; reading comprehension: $V = 15$, $p = 0.004$; speech comprehension: $V = 53$, $p = 0.142$; verbal production: $V = 78$, $p = 0.330$), and Cohen's *d* was used to quantify learning-dependent change in

326 hemispheric dominance for each modality. In both experiments, the same pattern of changes

327 emerged: large in reading comprehension, medium in speech comprehension, and small in verbal

328 production (Figure 4).

329   Finally, modality-wise clustering of joint L1-Ln lateralisation was plotted using 85% data

330 ellipses to examine overlap between modalities. One-way MANOVAs and a joint distribution

331 difference index were used to test and quantify the separation between: (i) comprehension (both

332 reading and speech) and production (Figure 5-I), and (ii) reading and speech comprehension

333 (Figure 5-II), and the effects of proficiency were tested using non-parametric two-sample/paired

334 tests of difference/change in cluster separation between modalities (Euclidean distance) and

335 quantified with percent change in the JDD. The one-way MANOVA modelled the joint L1-Ln

336 distribution differences between modalities, and the index quantified this difference by taking

337 into account the difference in both bivariate mean and spread of data, with values between 0

338 (overlapping distributions) and 1 (no similarities). MANOVAs revealed significant differences

339 between comprehension and production (Experiment I: basic proficiency group: $F_{(1.8,65.2)}=11.73$,

340 p=0.0005; advanced proficiency group: $F_{(1.9,63.8)}=22.96$, p=0.00000002; Experiment II: before

341 training: $F_{(1.7,73.2)}=21.67$, p=0.0000002, after training: $F_{(1.7,70.7)}=38.94$, p=0.0000000000004) and

342 with increasing proficiency, comprehension and production dissociated further (Experiment I:

343 the advanced proficiency group displayed 1042.35% greater comprehension-production

344 dissociation than the basic proficiency group, Mann-Whitney U test of group difference in cluster

345 separation: W = 67398, p = 0.000000000003; Experiment II: participants displayed 47.38% increase

346 in comprehension-production dissociation after training, Wilcoxon signed-rank test of post-

347 training change in cluster separation: V = 101769, p = 0.0000000000000002). There were no

348 significant differences in L1-Ln joint distribution between reading and speech comprehension

349 (Experiment I: basic proficiency group: $F_{(1.9,45.5)}=1.84$, p=0.18; advanced proficiency group:

350 $F_{(1.7,41.4)}=0.32$, p=0.71; Experiment II: Before Training: $F_{(1.9,71.2)}=1.98$, p=0.15, After Training:

351 $F_{(1.8, 60.5)}=2.09$, p=0.13), and reading and speech comprehension converged further with

352 increasing proficiency (Experiment I: the advanced group displayed 87.27% greater

353 comprehension-production overlap than the basic group, Mann-Whitney U test of group

354 difference in cluster separation: W = 18073, p = 0.177; Experiment II: participants displayed 27.13%

355 increase in comprehension-production overlap after training, Wilcoxon signed-rank test of post-

356 training change in cluster separation: V = 39306, p = 0.0005).

357

Discussion

359 In the present work, we examined hemispheric specialisation and learning-dependent

360 plasticity of the language network concurrently in three language systems: reading, speech

361 comprehension and verbal production. We conducted cross-sectional and longitudinal fMRI

362 experiments in separate populations of immersed language learners. Both experiment samples

363 had the same L1 (Spanish), but were contrasted in other factors: (i) early language experience:

364 monolingual vs sequential bilingual, (ii) language being learnt: Basque vs English, (iii)

365 phonological similarity with native language: high overlap vs low overlap, (iv) orthographic

366 depth: transparent vs opaque. Across these contrasting experimental designs and participant

367 groups, we found a highly consistent pattern of results in both experiments: (i) across native and

368 non-native languages, lateralisation for language comprehension was variable but language

369 production was strongly left-lateralised, and (ii) with increasing non-native language proficiency,

370 reading and speech comprehension displayed significant changes in hemispheric dominance

371 (reading > speech), while verbal production remained left-lateralised. The converging results

372 from separate experiments provide unique insight into the long-standing debate on hemispheric

373 specialisation of language and the effects of language experience (Gainotti, 1993; Price, 1998,

374 2012; Jung-Beeman, 2005; Hickok and Poeppel, 2007; Friederici, 2012; Hervé et al., 2013; Tzourio-

375 Mazoyer et al., 2016).

The first result showing variably-lateralised (bilateral at the group-level) comprehension and left-lateralised verbal production across different languages suggested that comprehension is flexible while verbal production is hard-wired to be left-lateralised. Previously, conflicting evidence from studies in monolinguals had led to a range of different conclusions and models of comprehension: from left-lateralised to partly-bilateral, bilateral, or right-lateralised function (Booth et al., 2000; Gaillard et al., 2000; Jung-Beeman, 2005; Hickok and Poeppel, 2007; Lidzba et al. 2011). Few studies have compared different modalities in the same participants, and though lateralisation was seen to be highly modality-dependent in the current study, it did not appear to depend on the exact task used, since lateralisation for the single-word overt tasks in the current study was consistent with results from far more complex discourse-level covert tasks in previous studies (Dehaene et al., 1997; Lidzba et al., 2011; Bhattasali et al., 2019). There were also subtle differences between the two experiments, with similar Ln lateralisation but differing central tendencies for L1 laterality. This pattern is consistent with the literature on the influence of early language experience: meta-analyses have found that early bilinguals (L2 acquired before age 6) typically show bilateral hemispheric involvement, while monolinguals and late bilinguals show greater left hemisphere dominance (Hull & Vaid, 2006, 2007; Bloch et al., 2009; Liu & Cao, 2016). Thus the convergent results in the present work indicate that inter-individual variability in lateralisation for language comprehension is not an artefact of task or methodology, but that instead, language comprehension is differently lateralised across individuals. Lesion studies in children have found dissociative effects of lesion side on comprehension and production: while lesions in the left hemisphere were associated with more severe delays in production compared to comprehension, comprehension delays were more common — but not universal — in children with right-hemisphere damage (Marchman, Miller and Bates, 1991; Thal et al., 1991). In line with these findings, developmental neuroimaging studies all found left-lateralised language production, but reached conflicting conclusions on comprehension, leading to a suggestion of differing maturational mechanisms for comprehension and production (Hervé et al., 2013).

Clinical studies have recommended that both comprehension and production tasks be used in determining language lateralisation for clinical purposes (Wilke et al., 2010; Lidzba et al., 2011; Vilasboas et al., 2017; Woodhead et al., 2018). Modality-dependent lateralisation, i.e. variably-lateralised comprehension vs left-lateralised production, could explain the long-standing conflicts among previous studies that used tasks tapping into different modalities, and shed new light on the question of functional specialisation for language.

Our analytic approach to examining learning-dependent changes in language lateralisation built on the observed inter-individual variability and used within-subject measures calculated with each subject's L1 as a baseline for their Ln. We used three measures — L1-Ln correlation, L1-Ln distance, and modality clustering — and quantified the change within each language system. These revealed that (i) L1 and Ln were similarly lateralised in lower-proficiency language learners and tended to dissociate with increasing Ln proficiency, (ii) the change was largest in reading, smaller in speech comprehension, and smallest in verbal production, and (iii) with increasing proficiency, comprehension and production dissociated, while reading and speech comprehension converged.

Convergence and dissociation of neural activation for different languages and language systems has been of considerable research interest. Neuroimaging studies of language have by and large come to the conclusion that all languages do indeed recruit the same language regions, and that language experience modulates the amount of overlap, leading to the "convergence hypothesis" (Perani and Abutalebi 2005; Gurunandan et al., 2019). The current study built on this finding, and characterised lateralisation patterns for L1 and Ln within the common language network, finding that increasing Ln-proficiency led to increasing dissociation in lateralisation between the two languages. There has been much debate on whether language control in bilinguals is language-specific or domain-general, with mixed evidence (Hernández et al., 2013), and it is possible that, apart from any changes in the involvement of language control regions, the greater hemispheric separation of languages in more proficient non-monolinguals also

428 contributes to their improved language control. Future studies looking concurrently at

429 dissociation within the language network and recruitment of language control regions are

430 needed to test this idea. Comprehension and production also dissociated with increasing Ln-

431 proficiency. In lower proficiency learners, there was lower separation between modalities,

432 possibly indicating variable strategies of Ln access and variable activation profiles (Dehaene et al.,

433 1997), but as individuals attained higher proficiency, their activation profiles stabilised and

434 became more uniform. Turning to the question of convergence between language systems, print-

435 speech convergence has been considered a universal signature of native language proficiency

436 (Shankweiler et al., 2008; Rueckl et al., 2015; Preston et al., 2016), but it is less well-studied in

437 multilinguals. In the current study, we found increasing convergence of joint L1-Ln lateralisation

438 for reading and speech comprehension with increasing language learning, suggesting that

439 reading-speech convergence is also sensitive to increasing Ln-proficiency.

440   The pattern of plasticity differences between the language systems, i.e. plasticity for

441 reading > speech comprehension > verbal production, was strikingly similar to their perceived

442 difficulty in real-world language learning in adults. Two observations support the idea that the

443 differential plasticity of language systems contribute to differential learning. First, learners in the

444 longitudinal study had switched languages from same to opposite hemispheres in reading within

445 a relatively short time-frame, while fewer did so for speech comprehension, and none for

446 production. Further, individuals who had L1 and Ln lateralised in opposite hemispheres

447 maintained this dissociation post-training, and individuals who had L1 and Ln in the same

448 hemisphere tended to dissociate post-training to varying degrees depending on the modality.

449 This suggested that opposite hemispheric dominance of languages could be advantageous for

450 language learning, and further, that shifts in hemispheric dominance are limited by the plasticity

451 of the sensory/motor cortices corresponding to each language system. Neuropsychological

452 evidence from stroke recovery patterns in adults who showed greater (but not complete)

453 recovery in comprehension than in production (Lomas and Kertesz 1978), as well as different

454 reorganisation patterns for comprehension and production (Musso et al., 1999; Heiss and Thiel

455 2006) further supports our conclusion. Though the visual, auditory and motor cortices are all

456 bilateral, each of them exhibits hemispheric advantages for processing specific features (Benke

457 and Kertesz, 1989; Deruelle and Fagot, 1997; Flinker et al., 2019; Albouy et al., 2020), and previous

458 studies with monolinguals have found differences in visual lateralisation of different writing

459 scripts (Tzeng et al., 1979; Kuo et al., 2001), asymmetrical sensitivity of the auditory cortices

460 (Friederici and Alter 2004; Boemio et al., 2005), and left-lateralisation of auditory and articulatory

461 motor areas (Morillon et al., 2010), pointing to differential potential for post-critical-period

462 plasticity of these sensory/motor regions that matches the converging pattern of language

463 system plasticity found in the current study. Second, the adolescent learners in the second

464 experiment displayed substantial neural changes after just three months of training, while the

465 adult learners in the first experiment displayed similar neural differences for a much larger

466 proficiency difference between groups. This finding is compatible with age-related decrease in

467 neural plasticity, and sheds further light on the source of the difficulty of late language learning.

468 However, despite the convergence of the neural results in experiments I and II, the modest

469 behavioural effect in Experiment II limited any further interpretation of the neural changes in

470 relation to behavioural outcomes at the individual level in naturalistic language learning. In sum,

471 taken together with previous evidence, the converging findings in the present work point to the

472 sensorimotor cortices playing a large role in both the lateralisation of language as well as the

473 asymmetric decrease in plasticity of the language network.

474 Methodological studies and reviews of language lateralisation have often warned against

475 over-interpretation of results from a single task, small regions of interest, or non-robust

476 analytical methods (e.g. Bradshaw et al., 2017a, 2017b; Bain et al., 2019). These were avoided in the

477 current study and interpretations were based on robust patterns of results verified by

478 corroborating analyses that were replicated in contrasting experiments. However, the current

479 study used classical single word tasks, and while the lateralisation results were consistent with

480   the findings from far more complex comprehension tasks (Dehaene et al., 1997; Lidzba et al., 2011;

481   Bhattasali et al., 2019), future studies are needed to establish whether the results presented in the

482   current study would be as or possibly even more pronounced in sentence/discourse processing

483   (Hagoort 2019). Further, a priori power analysis was not conducted nor was a replication sample

484   examined. The two experiments involved ecologically valid language learning, and the

485   lateralisation results were sensitive to participants' real-world language-learning progress, i.e.

486   CEFR level, rather than their performance or improvement on the in-scanner semantic tasks

487   involving high-frequency stimuli. In fact, while performance on the tasks was relatively uniform

488   across participants, lateralisation exhibited much larger variation in both languages, supporting

489   the idea of multi-factorial modulation of hemispheric specialisation (Tzourio-Mazoyer et al.,

490   2016), since participants were carefully selected to control for language backgrounds as much as

491   possible, but actual experimental control on early or previous language exposure was not

492   possible. The replication of findings in language learners at different levels of proficiency

493   suggested that the learning-dependent neural changes were not temporary, but further studies

494   are necessary to disentangle the effects of learning vs proficiency, and test lateralisation of

495   languages in early balanced bilinguals. Finally, the two experiments featured distinct L1-Ln

496   language pairs that were contrasted on factors such as overlap in language families, phonological,

497   and orthographic properties, but did not involve more sensory differences such as visual

498   differences between writing systems or auditory perception and motor production of tones, as in,

499   say, English-Chinese. Following from our idea that the sensorimotor cortices are the limiting

500   factor in language learning and its associated neural changes, it is possible that late acquisition of

501   a language that requires greater sensorimotor learning would entail smaller proficiency-

502   dependent neural changes in lateralisation and the size of the changes would decrease more

503   sharply with increasing age than in the current study.

504   In conclusion, our study design with cross-sectional and longitudinal experiments in

505   contrasting samples of real-world language learners, testing of different language systems, and a

506   multi-pronged analytical approach revealed robust and converging patterns of modality-

507   dependent lateralisation and plasticity of the language network. Our findings suggest that

508   language lateralisation for reading and speech comprehension is plastic well into adulthood,

509   while production shows strong left-hemisphere specialisation. Plasticity for reading was greater

510   than for speech comprehension which was in turn greater than for verbal production. Taken

511   together with previous evidence in the literature, we propose that hemispheric specialisation for

512   language may arise from the sensorimotor cortices, and that the differential plasticity of

513   language systems is tied to the plasticity of the associated sensorimotor systems.

References

Ahmad Z, Balsamo LM, Sachs BC, Xu B, Gaillard WD (2003) Auditory comprehension of language in young children: neural networks identified with fMRI. Neurology 60(10) 1598-1605

Albouy P, Benjamin L, Morillon B, Zatorre RJ (2020) Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. Science 367:1043-1047

Bain JS, Yeatman JD, Schurr R, Rokem A, Mezer AA (2019) Evaluating arcuate fasciculus laterality measurements across dataset and tractography pipelines. Hum Brain Mapp 40:3695-3711

Bates E (1993) Comprehension and production in early language development: Comments on Savage-Rumbaugh et al. Monogr Soc Res Child Dev 58:222-242

Bavelier D, Neville HJ (2002) Cross-modal plasticity: where and how? Nat Rev Neurosci 3:443-452

Benke T, Kertesz A (1989) Hemispheric mechanisms of motor speech. Aphasiology 3:627-641

Bhattasali S, Fabre M, Luh WM, Al Saied H, Constant M, Pallier C, Hale J (2019) Localising memory retrieval and syntactic composition: an fMRI study of naturalistic language comprehension. Lang Cogn Neurosci 34:491-510

Bloch C, Kaiser A, Kuenzli E, Zappatore D, Haller S, Franceschini R, Nitsch C (2009) The age of second language acquisition determines the variability in activation elicited by narration in three languages in Broca's and Wernicke's area. Neuropsychologia 47:625-633

Boatman D, Freeman J, Vining E, Pulsifer M, Miglioretti D, Minahan R, McKhann G (1999) Language recovery after left hemispherectomy in children with late-onset seizures. Ann Neurol 46:579-586

Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:389

Booth JR, MacWhinney B, Thulborn KR, Sacco K, Voyvodic JT, Feldman HM (2000) Developmental and lesion effects in brain activation during sentence comprehension and mental rotation. Dev Neuropsychol 18(2) 139-169

Bradshaw AR, Bishop DV, Woodhead ZV (2017) Methodological considerations in assessment of language lateralisation with fMRI: a systematic review. PeerJ 5:e3557

Bradshaw AR, Thompson PA, Wilson AC, Bishop DV, Woodhead ZV (2017) Measuring language lateralisation with different language tasks: a systematic review. PeerJ 5:e3929

Broca P (1863) Localisation des fonctions cérébrales: Siége de langage articulé. Bulletins de la Société d'Anthropologie de Paris 4:200-208

Dale AM (1999) Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109-114

Dax G (1863) Observations tendant à prouver la coïncidence constante des dérangements de la parole avec une lésion de l'hémisphère gauche du cerveau. C R Acad Sci III 56:536

Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002) Functional neuroimaging of speech perception in infants. Science 298:2013-2015

Dehaene S, Dupoux E, Mehler J, Cohen L, Paulesu E, Perani D, Le Bihan D (1997) Anatomical variability in the cortical representation of first and second language. Neuroreport 8:3809-3815

Deruelle C, Fagot J (1997) Hemispheric lateralisation and global precedence effects in the processing of visual stimuli by humans and baboons (Papio papio). Laterality 2:233-246

Duffau H, Capelle L, Denvil D, Sichez N, Gatignol P, Lopes M, Van Effenterre R (2003) Functional recovery after surgical resection of low grade gliomas in eloquent brain: hypothesis of brain compensation. J Neurol Neurosurg Psychiatry 74:901-907

Duffau H, Denvil D, Capelle L (2002) Long term reshaping of language sensory and motor maps after glioma resection: a new parameter to integrate in the surgical strategy. J Neurol Neurosurg Psychiatry 72:511-516

Enge A, Friederici AD, Skeide MA (2020) A meta-analysis of fMRI studies of language comprehension in children. NeuroImage 116858

Flinker A, Doyle WK, Mehta AD, Devinsky O, Poeppel D (2019) Spectrotemporal modulation provides a unifying framework for auditory cortical asymmetries. Nat Hum Behav 3:393-405

Friederici AD (2012) The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn Sci 16:262-268

Friederici AD, Alter K (2004) Lateralization of auditory language functions: a dynamic dual pathway model. Brain Lang 89:267-276

Gaillard WD, Sachs BC, Whitnah JR, Ahmad Z, Balsamo LM, Petrella JR, Grandin CB (2003) Developmental aspects of language processing: fMRI of verbal fluency in children and adults. Hum Brain Mapp 18:176-185

Gainotti G (1993) The riddle of the right hemisphere's contribution to the recovery of language. Eur J Disord Commun 28:227-246

Gurunandan K, Carreiras M, Paz-Alonso PM (2019) Functional plasticity associated with language learning in adults. NeuroImage 201:116040

Hagoort P (2013) MUC (memory unification control) and beyond. Front Psychol 4:416

Hagoort P (2019) The neurobiology of language beyond single-word processing. Science 366:55-58

Heiss WD, Thiel A (2006) A proposed regional hierarchy in recovery of post-stroke aphasia. Brain Lang 98:118-123

Hernández M, Martin CD, Barceló F, Costa A (2013) Where is the bilingual advantage in task-switching? J Mem Lang 69:257-276

Hervé PY, Zago L, Petit L, Mazoyer B, Tzourio-Mazoyer N (2013) Revisiting human hemispheric specialization with neuroimaging. Trends Cogn Sci 17:69-80

Hickok G, Poeppel D (2007) The cortical organization of speech processing. Nat Rev Neurosci 8:393

Hope TM, Leff AP, Prejawa S, Bruce R, Haigh Z, Lim L, Seghier ML (2017) Right hemisphere structural adaptation and changing language skills years after left hemisphere stroke. Brain 140:1718-1728

Horowitz-Kraus T, Grainger M, DiFrancesco M, Vannest J, Holland SK, CMIND Authorship Consortium (2015) Right is not always wrong: DTI and fMRI evidence for the reliance of reading comprehension on language-comprehension networks in the right hemisphere. Brain Imaging Behav 9:19-31

Hull R, Vaid J (2006) Laterality and language experience. Laterality 11:436-464

Hull R, Vaid J (2007) Bilingual language lateralization: A meta-analytic tale of two hemispheres. Neuropsychologia 45:1987-2008

Jung-Beeman M (2005) Bilateral brain processes for comprehending natural language. Trends Cogn Sci 9:512-518

Kaplan EF, Goodglass H, Weintraub S (1983) The Boston naming test. Philadelphia, Lea Febiger

Kleim JA, Jones TA (2008) Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 51:S225-S239

Kovelman I, Shalinsky MH, Berens MS, Petitto LA (2008) Shining new light on the brain's "bilingual signature": a functional Near Infrared Spectroscopy investigation of semantic processing. Neuroimage 39:1457-1471

Kuo WJ, Yeh TC, Duann JR, Wu YT, Ho LT, Hung D, Hsieh JC (2001) A left-lateralized network for reading Chinese words: a 3 T fMRI study. Neuroreport 12:3997-4001

Lidzba K, Schwilling E, Grodd W, Krägeloh-Mann I, Wilke M (2011) Language comprehension vs language production: age effects on fMRI activation. Brain Lang 119:6-15

Lomas J, Kertesz A (1978) Patterns of spontaneous recovery in aphasic groups: A study of adult stroke patients. Brain Lang 5:388-401

Marchman VA, Miller R, Bates EA (1991) Babble and first words in children with focal brain injury Appl Psycholinguist 12:1-22

Mårtensson J, Eriksson J, Bodammer NC, Lindgren M, Johansson M, Nyberg L, Lövdén M (2012) Growth of language-related brain areas after foreign language learning. NeuroImage 63:240-244

Mazaika PK, Hoeft F, Glover GH, Reiss AL (2009) Methods and software for fMRI analysis of clinical subjects. Neuroimage 47:S58

Morillon B, Lehongre K, Frackowiak RS, Ducorps A, Kleinschmidt A, Poeppel D, Giraud A L (2010) Neurophysiological origin of human brain asymmetry for speech and language. Proc Natl Acad Sci U S A 107:18688-18693

Musso M, Weiller C, Kiebel S, Müller SP, Bülau P, Rijntjes M (1999) Training-induced brain plasticity in aphasia. Brain 122:1781-1790

Papanicolaou AC, Moore BD, Levin HS, Eisenberg HM (1987) Evoked potential correlates of right hemisphere involvement in language recovery following stroke. Arch Neurol 44:521-524

Park HR, Badzakova-Trajkov G, Waldie KE (2012) Language lateralisation in late proficient bilinguals: A lexical decision fMRI study. Neuropsychologia 50:688-695

Payne BR, Lomber SG (2001) Reconstructing functional systems after lesions of cerebral cortex. Nat Rev Neurosci 2:911-919

Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE (2011) Statistical Parametric Mapping: the Analysis of Functional Brain Images. Elsevier

Perani D, Abutalebi J (2005) The neural basis of first and second language processing. Curr Opin Neurobiol 15:202-206

Perani D, Saccuman M C Scifo P Anwander A Spada D Baldoli C Friederici A D (2011) Neural language networks at birth. Proc Natl Acad Sci U S A 108:16056-16061

Preston JL, Molfese PJ, Frost SJ, Mencl WE, Fulbright RK, Hoeft F, Landi N, Shankweiler D, Pugh KR (2016) Print-speech convergence predicts future reading outcomes in early readers. Psychol Sci 27:75–84

Price C J (1998) The functional anatomy of word comprehension and production. Trends Cogn Sci 2:281-288

Price C J (2012) A review and synthesis of the first 20 years of PET and fMRI studies of heard speech spoken language and reading. Neuroimage 62:816-847

Rasmussen T and Milner B (1975) Clinical and surgical studies of the cerebral speech areas in man. In: Cerebral localization, pp238-257. Springer: Berlin Heidelberg.

Rueckl JG et al. (2015) Universal brain signature of proficient reading: Evidence from four contrasting languages. Proc Natl Acad Sci U S A 112:15510-15515.

Schlegel AA, Rudelson JJ, Tse PU (2012) White matter structure changes as adults learn a second language. J Cogn Neurosci 24:1664-1670

Shankweiler D, Mencl WE, Braze D, Tabor W, Pugh KR, Fulbright RK (2008) Reading differences and brain: cortical integration of speech and print in sentence processing varies with reader skill. Dev Neuropsychol 33:745–775

Skeide MA, Friederici AD (2016) The ontogeny of the cortical language network. Nat Rev Neurosci 17:323-332

Szaflarski JP, Holland SK, Schmithorst VJ, Byars AW (2006) fMRI study of language lateralization in children and adults. Hum Brain Mapp 27:202-212

Thal DJ, Marchman V, Stiles J, Aram D, Trauner D, Nass R, Bates E (1991) Early lexical development in children with focal brain injury. Brain Lang 40:491-527

Tzeng OJ, Hung DL, Cotton B, Wang WS (1979) Visual lateralisation effect in reading Chinese characters. Nature 282:499

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single- subject brain. NeuroImage 15:273-289

Tzourio-Mazoyer N, Perrone-Bertolotti M, Jobard G, Mazoyer B, Baciu M (2016) Multi-factorial modulation of hemispheric specialization and plasticity for language in healthy and pathological conditions: A review, Cortex 86:314-339

Vigneau M, Beaucousin V, Hervé P Y, Jobard G, Petit L, Crivello F, Tzourio-Mazoyer N (2011) What is right-hemisphere contribution to phonological lexico-semantic and sentence processing?: Insights from a meta-analysis. Neuroimage 54:577-593

Vilasboas T, Herbet G, Duffau H (2017) Challenging the myth of right nondominant hemisphere: lessons from corticosubcortical stimulation mapping in awake surgery and surgical implications. World Neurosurg 103:449-456

Vingerhoets G, Van Borsel J, Tesink C, Van den Noort M, Deblaere K, Seurinck R, Achten E (2003) Multilingualism: an fMRI study. NeuroImage 20:2181-2196

Wilke M, Lidzba K (2007) LI-tool: a new toolbox to assess lateralization in functional MR-data J Neurosci Methods 163:128–136

Wilke M, Pieper T, Lindner K, Dushe T, Holthausen H, Krägeloh-Mann I (2010) Why one task is not enough: functional MRI for atypical language organization in two children. Eur J Paediatr Neurol 14:474-478

Woodhead ZV, Rutherford HA, Bishop DV (2018) Measurement of language laterality using functional transcranial Doppler ultrasound: a comparison of different tasks. Wellcome Open Res 3

Xiang H, van Leeuwen TM, Dediu D, Roberts L, Norris DG, Hagoort P (2015) L2-proficiency-dependent laterality shift in structural connectivity of brain language pathways. Brain Connect 5:349-361

Tables and Legends

| Table 1. Participant demographics and linguistic scores						
	Experiment I			Experiment II		
	Basic proficiency group	Advanced proficiency group	Statistical tests	Before training	After training	Statistical tests
Age	42.9 (10.1)	44.5 (10.5)	t(26.9) = 0.44, p = 0.66, two-sample t-test	17.2 (0.6)		-
Gender	7 female, 7 male	8 female, 7 male	χ^2 (1) = 0, p = 1 chi-square test for independence	16 female, 3 male		-
Proficiency						
L1	99.35 (1.88)	99.64 (0.77)	t(13.9) = 0.52, p = 0.61, two-sample t-test	99.11 (1.49)	99.26 (1.15)	t(23) = 0.90, p = 0.56, paired t-test
Ln	52.6 (14.66)	87.96 (10.58)	t(19.4) = -7.02, p = 0.0000009, two-sample t-test, Cohen's d = 2.82	58.00 (11.73)	62.89 (12.82)	t(23) = 2.98, p = 0.006, paired t-test, Cohen's d = 0.42
Note: Values correspond to the mean with standard deviation in parentheses.						

Figure 1. Behavioural measures Accuracy and Response Time for in-scanner semantic tasks plotted as a function of Group, Language, and Modality in Experiment I (A), and as a function of Training, Language, and Modality in Experiment II (B). Error bars represent standard deviation and asterisks statistically significant differences (*** p < 0.001, ** p<0.01, * p<0.05).

Figure 2. Laterality indices plotted as a function of Group, Language, and Modality in Experiment I (A), and Training, Language, and Modality in Experiment II (B). Laterality indices were obtained from individual whole brain activation in the neuroanatomical language network, and the respective line graphs display mean and standard deviation of laterality indices across participants in each Modality and Language.

Figure 3. Linear associations between L1 and Ln lateralisation indices (Pearson´s *r*) as a function of Group and Modality in Experiment I (A) and of Training and Modality in Experiment II (B). Cohen's *q* quantified the learning-dependent changes in L1-Ln correlation in each Modality.

Figure 4. Lateralised dissociation indices (LDI) as a function of Group and Modality in Experiment I (A), and Training and Modality in Experiment II (B). Positive values indicate that L1 and Ln lateralised to opposite hemispheres, and negative values indicate that L1 and Ln lateralised to the same hemisphere. Cohen's *d* quantified the learning-dependent changes in LDI in each Modality.

Figure 5. Modality-wise clustering (I: comprehension versus production, II: reading versus speech comprehension) in joint distributions of L1-Ln lateralisation indices plotted as a function of Group in Experiment I (A) and as a function of Training in Experiment II (B). A Joint Distribution Difference index with values between 0 and 1 quantified overlap in each group, with higher values indicating larger separation between modalities. Asterisks represent statistically significant differences (p < 0.001).
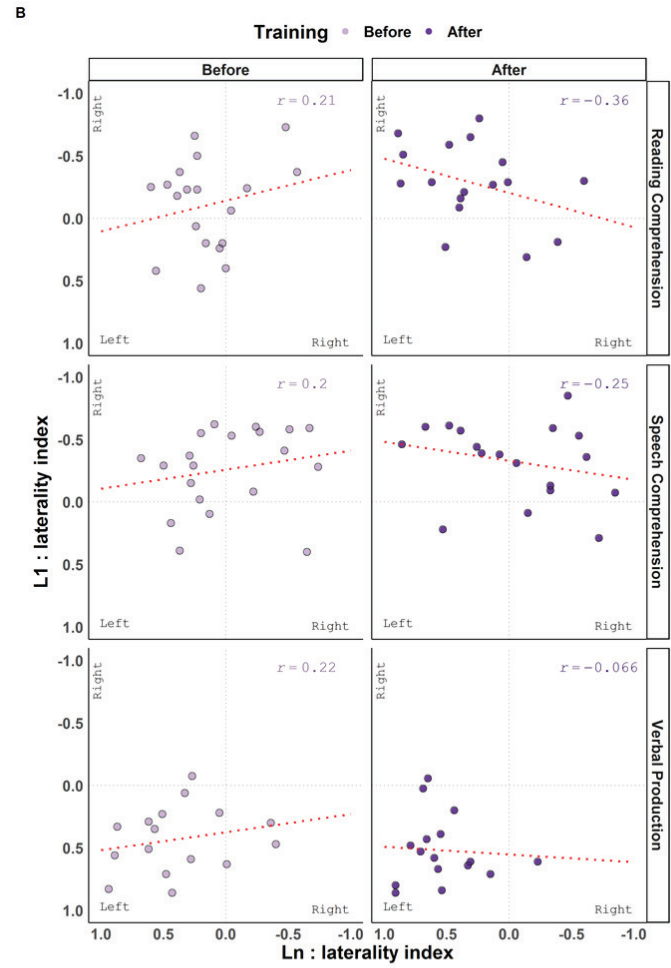
A

Group ● Basic ● Advanced
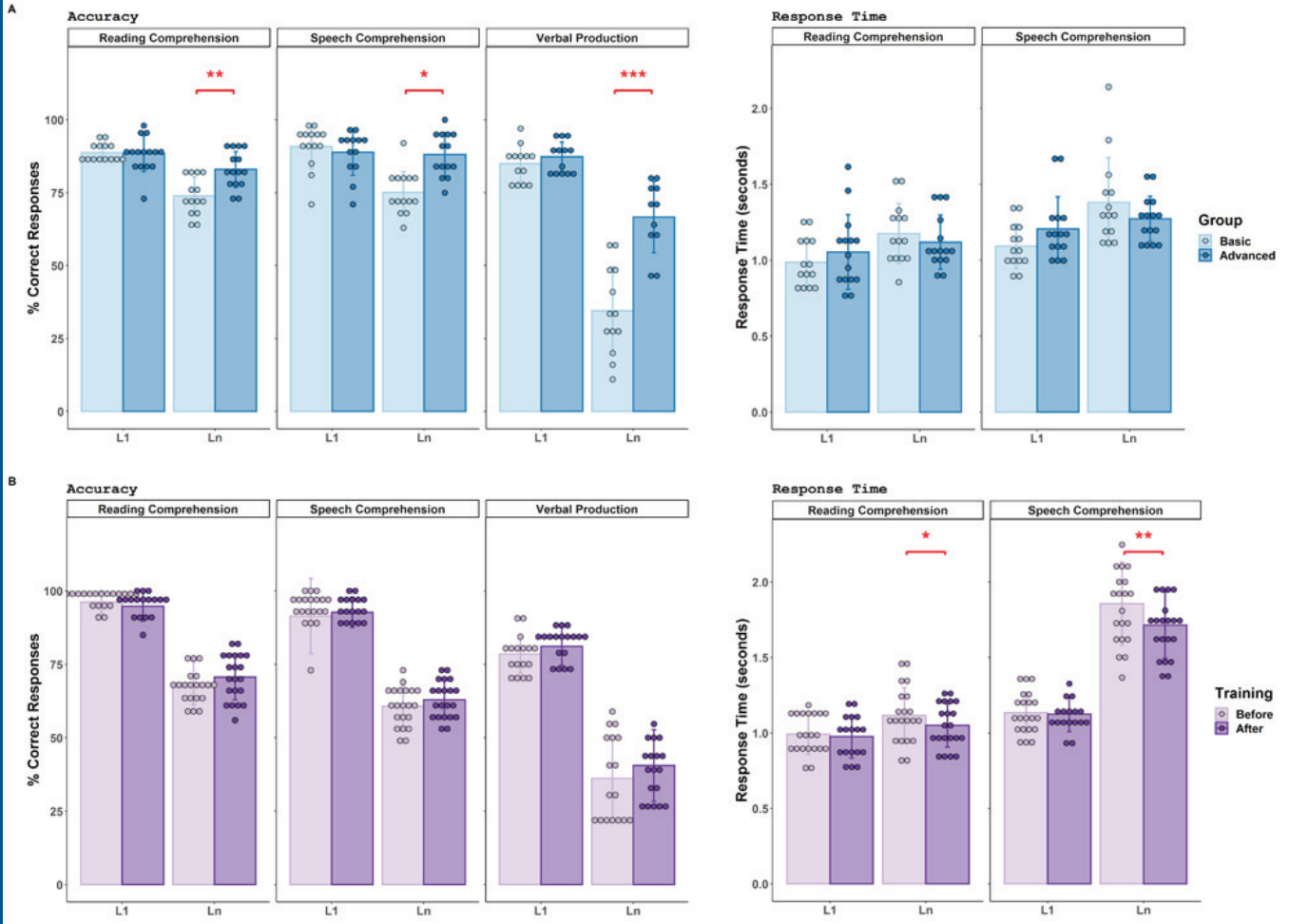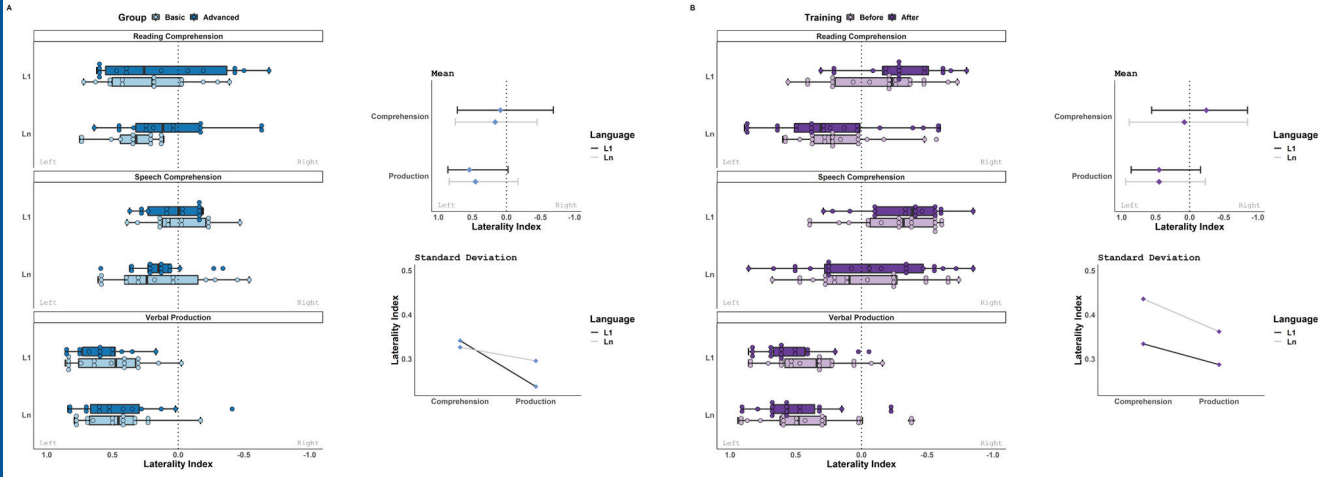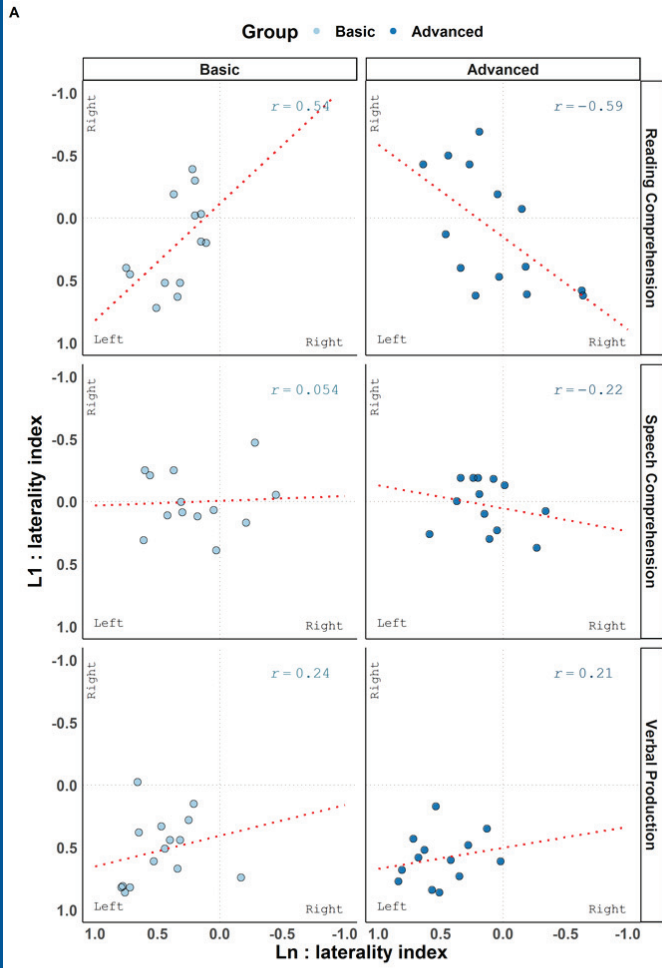
Task	Cohen's q	Effect
Reading Comprehension	1.282	Large
Speech Comprehension	0.278	Small
Verbal Production	0.032	Negligible

B

Training ● Before ● After

Task	Cohen's q	Effect
Reading Comprehension	0.59	Large
Speech Comprehension	0.458	Medium
Verbal Production	0.29	Small

I (A)

Modality ■ Comprehension ■ Production

Basic — Difference = 0.048***
Advanced — Difference = 0.549***

I (B)

Modality ■ Comprehension ■ Production

Before — Difference = 0.190***
After — Difference = 0.280***

II (A)

Modality ■ Reading Comprehension ■ Speech Comprehension

Basic — Difference = 0.175
Advanced — Difference = 0.022

II (B)

Modality ■ Reading Comprehension ■ Speech Comprehension

Before — Difference = 0.066
After — Difference = 0.048

L1 : laterality index
Ln : laterality index