

Neuroimaging of the bilingual brain: Structural brain correlates of listening and speaking in a second language



Patricia K. Kuhl^{a,*}, Jeff Stevenson^a, Neva M. Corrigan^{a,b}, Jasper J.F. van den Bosch^a, Dilara Deniz Can^a, Todd Richards^{a,b}

^a Institute for Learning & Brain Sciences, University of Washington, Seattle, WA 98195, USA

^b Department of Radiology, University of Washington, Seattle, WA 98195, USA

ARTICLE INFO

Article history:

Received 8 January 2016

Revised 12 July 2016

Accepted 12 July 2016

Available online 1 August 2016

Keywords:

Diffusion tensor imaging

Bilingual

Language immersion

Fractional anisotropy

White matter

ABSTRACT

Diffusion tensor imaging was used to compare white matter structure between American monolingual and Spanish-English bilingual adults living in the United States. In the bilingual group, relationships between white matter structure and naturalistic immersive experience in listening to and speaking English were additionally explored. White matter structural differences between groups were found to be bilateral and widespread. In the bilingual group, experience in *listening* to English was more robustly correlated with decreases in radial and mean diffusivity in anterior white matter regions of the left hemisphere, whereas experience in *speaking* English was more robustly correlated with increases in fractional anisotropy in more posterior left hemisphere white matter regions. The findings suggest that (a) foreign language immersion induces neuroplasticity in the adult brain, (b) the degree of alteration is proportional to language experience, and (c) the modes of immersive language experience have more robust effects on different brain regions and on different structural features.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Converging evidence suggests that bilingualism is associated with brain structural alterations (for a review, see [García-Pentón, García, Costello, Duñabeitia, & Carrieras, 2016](#)). There is also evidence that the quality and quantity of second-language (L2) learning experience affects brain structure ([Stein, Winkler, Kaiser, & Dierks, 2014](#)). However, there has been little investigation of possible differential effects of language mode on brain structural characteristics associated with L2 learning. A well established characterization of regional specialization for receptive and expressive language in monolinguals exists, where the left hemisphere has long been recognized as generally dominant in language, and specific brain regions (e.g. Broca's and Wernicke's areas) have historically been identified as specialized for specific aspects of speech production and comprehension ([Kuhl & Damasio, 2013](#)). Moreover, contemporary models of monolingual language processing feature “dual streams” of information processing, with different streams for receptive and expressive lan-

guage ([Hickok & Poeppel, 2004, 2007](#)). In these models, the streams originate from areas involved in the early processing of speech information in the general area of the superior temporal gyrus. A ventral stream, projecting to the middle and inferior temporal cortices, processes the conceptual meaning of phonological information and is involved in listening to and understanding speech. A dorsal stream, projecting towards inferior parietal and posterior frontal lobe regions, plays a role in integrating auditory information with motor system representations, and allows for the production of speech sounds based on previous auditory input ([Saur et al., 2008](#)).

Recent evidence from the neuroimaging literature indicates that there may be a similar regional specialization for the two modes of language in second language learning, and that connections between regions involved in perception and production may play a crucial role. For example, a number of studies indicate that listening to speech results in the generation of internal motor models of speech production. In adults, investigators have reported fMRI activation of areas involved with speech production while listening to non-native phonetic contrasts ([Callan, Jones, Callan, & Akahane-Yamada, 2004](#); [Callan et al., 2003](#)), as well as during audiovisual native-language speech perception ([Skipper, van Wassenhove, Nusbaum, & Small, 2007](#)). Previous work from our laboratory indicates that for young children who are in the

* Corresponding author at: Institute for Learning & Brain Sciences, Box 357988, University of Washington, Seattle, WA 98195, USA.

E-mail addresses: pkkuhl@uw.edu (P.K. Kuhl), mrjeffs@uw.edu (J. Stevenson), neva@uw.edu (N.M. Corrigan), jasperb@uw.edu (J.J.F. van den Bosch), dilara.deniz@gmail.com (D.D. Can), toddr@uw.edu (T. Richards).

sensitive period for phonetic learning, listening to speech results in activation of areas involved in motor speech planning, including Broca's area and the cerebellum (Kuhl, Ramirez, Bosseler, Lin, & Imada, 2014).

Diffusion tensor imaging (DTI), a neuroimaging tool for measuring microscopic characteristics of white matter, has been used to provide evidence of differences between the bilingual and monolingual brain. DTI measures are sensitive to numerous white matter structural features, including degree of myelination, axonal caliber, fiber packing density, intra-cellular volume fraction of extra neurite glial cells and proportion of crossing fibers. The DTI measure most commonly reported is fractional anisotropy (FA), which measures the directional asymmetry of the diffusion of water molecules. White matter FA values are thought to be influenced by characteristics such as axonal density, degree of myelination, and geometrical fiber arrangement (Assaf & Pasternak, 2008). Significant differences in FA between monolinguals and bilinguals have been reported for child (Mohades et al., 2012, 2015), adult (Cummine & Boliek, 2012; Pliatsikas, Moschopoulou, & Saddy, 2015), and aging populations (Gold, Johnson, & Powell, 2013; Luk, Bialystok, Craik, & Grady, 2011), most consistently in the corpus callosum (CC), inferior fronto-occipital fasciculus (IFOF), and the superior longitudinal fasciculus (SLF). However, there is considerable variability in findings across studies, particularly in the direction of the FA differences (Table 1). Three studies have reported regional *increases* in FA in bilinguals as opposed to monolinguals (Luk et al., 2011; Mohades et al., 2015; Pliatsikas et al., 2015), two studies have reported regional *decreases* in FA in bilinguals as compared to monolinguals (Cummine & Boliek, 2012; Gold et al., 2013), and one study (Mohades et al., 2012) has reported both patterns – regional increases in FA in bilinguals as compared to monolinguals in the left IFOF and regional decreases in FA in bilinguals as compared to monolinguals in anterior corpus callosum fibers that project to the orbitofrontal cortex.

The objectives of the current study are: (1) to characterize differences in white matter structure between monolingual and Spanish-English bilingual adults living in the United States, and (2) to explore the effects of second language experience, particularly in *listening* to as compared to *speaking* English, on white matter structural features. We utilized an expanded set of DTI measures to characterize white matter structure, including not only FA, but also radial diffusivity (RD), a measure of diffusion of water molecules perpendicular to white matter tracts; axial diffusivity (AD), a measure of diffusion of water molecules parallel to white matter tracts; and mean diffusivity (MD), an average of displacement in all measured directions. We also examined the mode of anisotropy (MO), which provides information about the three-dimensional shape of the diffusion tensor, ranging from planar (indicating a large proportion of crossing fibers) to linear (indicating a single fiber orientation).

2. Materials and methods

2.1. Participants

Thirty-one adults, 16 Spanish-English speaking bilinguals (8 females) and 15 English-speaking monolinguals (9 females), participated in this study. Mean age at participation was 26.9 for the bilingual group (range = 22–32 years) and 25.5 for the monolingual group (range = 21–33 years). Participants reported no speech, language or hearing difficulties. All monolingual participants were born in the US and were native English language speakers. Bilingual participants were primarily individuals who immigrated to the United States in adulthood, with the exception of one participant who was born in the United States. All participants were

asked to complete a demographic survey that included questions on lifetime language use history. Participants were additionally asked to complete a language experience and usage questionnaire that documented duration of residence in the US, language use in different social contexts, as well as language exposure by medium. This questionnaire included self-reports of English comprehension and speaking proficiency (on a five-point Likert scale) as well as listening and speaking experience for English and Spanish in 3 year increments since birth (Garcia-Sierra, 2007; Garcia-Sierra, Diehl, & Champlin, 2009; Garcia-Sierra, Ramirez-Esparza, Silva-Pereyra, Siard, & Champlin, 2012). Socioeconomic status (SES) scores for all study participants were calculated based on occupation and education using the Hollingshead Index (Hollingshead, 2011). Hollingshead metrics were additionally used to calculate scores for education alone. Detailed participant information is provided in Supplementary Table S1. No significant differences between education level (Mann-Whitney $U = 107.0$, $p = 0.570$, two-tailed) or SES (Mann-Whitney $U = 117.0$, $p = 0.901$, two-tailed) at the time of data collection were found between the monolingual and bilingual groups. There were also no significant group differences in age (two-tailed t -test $p = 0.274$). The project was approved by the University of Washington Institutional Review Board, and written informed consent was obtained from all study participants.

2.2. DTI acquisition methods

DTI data were acquired on a Philips (Netherlands) 3T Achieva scanner using an 8-channel Philips head coil using a standard spin-echo/echo-planar DTI pulse sequence with scan parameters: TR/TE 7520/70 ms, voxel size $1.7 \times 1.7 \times 2$ mm (5.78 mm³), b value of 1000, 32 diffusion-encoded directions, 1 average, and acceleration factor of 2.3.

2.3. DTI data processing

Analysis of the DTI data was carried out using TBSS (Tract-Based Spatial Statistics; Smith et al., 2006), a part of the FSL Software Library (Smith et al., 2004). In TBSS, FA images are created by fitting a tensor model to the raw diffusion data using the FDT tool, and then brain-extracted using BET (Smith, 2002). Subjects' FA data are then aligned into a common space using the nonlinear registration tool FNIRT (Andersson, Jenkinson, & Smith, 2007a, 2007b), which uses a b-spline representation of the registration warp field (Rueckert et al., 1999). Next, a mean FA image is created and thinned to create a mean FA skeleton, which represents the centers of all tracts common to the group. Each subject's aligned FA data is then projected onto this skeleton. We additionally used the warp field to project MD, AD, RD, and MO onto the skeleton. The data were filtered for gender effects using FSL's Regfilt tool.

2.4. Group comparisons

Voxel-wise statistical comparisons were performed with FSL's Randomise tool (Winkler, Ridgway, Webster, Smith, & Nichols, 2014), using the "Threshold-Free Cluster Enhancement" (TFCE) option, with five thousand two-tailed Monte Carlo permutations. A higher level design matrix with an unpaired two group design was specified in FSL's FEAT tool. Twenty-one white matter tracts in the cerebrum were investigated in our analysis, as defined by the ICBM-DTI-81 white matter labels atlas, and the Johns Hopkins University white matter tractography atlas (JHU-ICBM-tracts-maxprob; Hua et al., 2008; Mori, Wakana, Nagae-Poetscher, & van Zijl, 2005; Wakana et al., 2007), which are both available in FSL. The JHU atlas was modified to correct for minor imperfections in the maxprob 0 threshold digital file and to add the middle corpus callosum region. The cerebellar regions were defined by

Table 1
DTI fractional anisotropy (FA) findings of white matter structural differences between monolinguals and bilinguals.

Author	Control Country of Origin	Control Language	Bilingual Country of Origin	L1	L2	N (Bi/Mono)	AoA (years)	Age in years mean (sd)	FA findings	White Matter Tracts
Luk et al. (2011)	Canada England Trinidad	English	Canada France Hungary Estonia Zimbabwe Morocco Romania Germany Poland	French Hungarian Hebrew Estonian Shona Yiddish Bulgarian Russian Ukrainian	English	14/14	<11	Mono 70.6 (3.1) Bi 70.3 (3.8)	Bi > Mono	R IFOF L&R SLF R UF CC
Mohades et al. (2012)	Europe	French or Dutch	NS	French or Dutch	Various Romance or Germanic languages	15 Sim Bi 15 Seq Bi 10 Mono	Sim Bi from birth Seq Bi > 3	Sim Bi 9.3 (0.9) Seq Bi 9.7 (0.9) Mono 9.6 (1.0)	Sim Bi > Mono Sim Bi > Seq Bi Mono > Sim Bi	L IFOF AC-OL
Cummine and Boliek (2012)	Canada	English	NS	Chinese	English	12/11	>5	Mono 28.5 (8.7) Bi 24.2 (4.1)	Mono > Bi	R IFOF R SUP ATR L&R INF ATR
Gold et al. (2013)	USA	English	NS	English	Igbo Swahili Filipino French German Gujarati Hindi Konkani Spanish	20/63	≤10	Mono 64.4 (5.1) Bi 63.9 (4.0)	Mono > Bi	L&R ILF/IFOF FORNIX CC
Mohades et al. (2015)	Europe	French or Dutch	NS	French or Dutch	Various Romance or Germanic languages	14 Sim Bi 16 Seq Bi 10 Mono	Sim Bi from birth Seq Bi > 3	Sim Bi 11.4 (0.9) Seq Bi 11.3 (1.0) Mono 11.1 (0.8)	Sim Bi > Mono	L IFOF
Pliatsikas et al. (2015)	Europe	English	Greece Italy China Poland Serbia France Hong Kong Bulgaria Netherlands Romania Germany Spain Uganda	Greek Italian Mandarin Polish Serbian French Cantonese Bulgarian Dutch Romanian German Spanish	English	20/25	Mean 10 SD 4.2	Mono 28.2 (5.3) Bi 31.9 (8.1)	Bi > Mono	CC L&R IFOF L&R SLF L&R UF

Mono: monolingual; Sim Bi: simultaneous bilingual; Seq Bi: sequential bilingual; Bi: bilingual; IFOF: inferior fronto-occipital fasciculus; SLF: superior longitudinal fasciculus; UF: uncinate fasciculus; CC: corpus callosum; AC-OL: anterior corpus callosum with projections to orbitofrontal cortex; ATR: anterior thalamic radiations; ILF: inferior longitudinal fasciculus; NS: not specified.

combining all cerebellar structures in the JHU MNI SS WMPM atlas (Oishi, Faria, & Mori, 2010) in each hemisphere. A significance threshold of $p = 0.05$ was utilized and all p -values were corrected for multiple comparisons. Group statistical maps were co-registered to these atlases, and a regional analysis was performed to identify clusters with significantly different FA, RD, AD, MD, or MO between groups within each tract and cerebellar hemisphere. Custom software was written to calculate average p -values, voxel counts, and average DTI measures for all significant voxels within each white matter region for each group.

2.5. Correlations with language immersion measures

For each bilingual subject, average reported experience in listening to English and in speaking English while residing in the

US were multiplied by the duration of residence in the US to yield measures reflecting English language experience in both of these modes during language immersion. A log transformation was applied to these estimates in order to address two issues. First, the duration of residence in the US was not normally distributed. Second, we expected that relationships between the behavioral and white matter structural measures would be nonlinear due to biological constraints. For example, degree of axonal myelination or fiber packing density cannot increase indefinitely. Significant correlations between these language measures and white matter structural measures were identified using the FSL Randomise tool combined with FSL's General Linear Model (GLM) tool (www.fmrib.ox.ac.uk/fsl/fsl/list) which generates p -value maps. All p -values were corrected for multiple comparisons. The data were filtered for gender effects using FSL's Regfilt tool. Five thousand two-

tailed Monte Carlo permutation tests were run to calculate correlations between language immersion measures and FA, RD, AD, MD, and MO. The threshold-free cluster enhancement (TFCE) option was utilized for the automatic identification of clusters. An identical analysis was performed for both hemispheres of the cerebellum. Custom software was written to generate average Pearson correlation coefficients, *p*-values, voxel counts, and DTI measures for all voxels showing significant correlations in each white matter region. File versions and attributes in the analysis pipeline were tracked using NiproV (van den Bosch & Tavabi, 2016).

3. Results

3.1. Group differences

Group differences between DTI measures for the monolingual and bilingual samples in both the cerebrum and the cerebellum were diffuse and bilateral. Significantly higher FA, and significantly lower RD and MD, were found in monolinguals as compared to bilinguals in clusters of voxels located along all of the 21 cerebral white matter tracts studied, as well in both cerebellar hemispheres. Mean DTI measures showing significant group differences in measures of white matter structure by tract are shown in Table 2. No significant differences in AD between groups were found. In addition to the findings shown in this table, MO in the same clusters exhibiting significantly lower FA was more negative in the bilingual sample than the monolingual sample, and MO was found to be significantly different between groups for one cluster of voxels located in the left SLF. Detailed MNI coordinate and voxel count

information is provided in Table S2 of the supplementary materials.

3.2. Correlations between duration of residence in US and white matter structure

Significant correlations between the logarithm of the duration of residence (IRes) in the United States and DTI measures in the cerebrum and cerebellum in bilinguals were widespread and bilaterally distributed (Fig. S1 of the supplementary materials, blue regions). In the cerebrum, significant correlations between FA and IRes were positive and more predominant in the left hemisphere. Significant correlations between IRes and RD, AD, and MD were negative, and found bilaterally in clusters located in all 21 tracts assessed. Table 3 shows significant Pearson correlation coefficients (*p* < 0.05, corrected for multiple comparisons and gender) for each tract. Detailed MNI coordinate and voxel count information is provided in Table S3 of the supplementary materials, along with representative scatter plots of correlations between IRes and DTI measures of FA and RD for the left IFOF (Fig. S2 of the supplementary materials). In the cerebellum, significant negative correlations were found between IRes and RD, MD, and AD in the left and right cerebellar hemispheres. No significant correlations were found between IRes and FA in the cerebellum.

3.3. Correlations between measures of English speaking and listening in the US and white matter structure

When duration of residence in the US was adjusted for reported amount of time speaking English (IExpSp) and listening to English

Table 2
DTI measures for voxels showing significant differences between groups (*p* < 0.05) by tract: mean (std. error).

Tract (by associated function) ^a	Fractional Anisotropy		Radial Diffusivity (mm ² /s)		Mean Diffusivity (mm ² /s)	
	Bilingual	Monolingual	Bilingual	Monolingual	Bilingual	Monolingual
<i>Executive</i>						
ATR L	0.41 (0.004)	0.44 (0.004)	5.3e ⁻⁴ (8e ⁻⁶)	4.9e ⁻⁴ (6e ⁻⁶)	7.0e ⁻⁴ (9e ⁻⁶)	6.5e ⁻⁴ (7e ⁻⁶)
ATR R	0.32 (0.004)	0.34 (0.006)	5.5e ⁻⁴ (8e ⁻⁶)	5.2e ⁻⁴ (8e ⁻⁶)	6.9e ⁻⁴ (8e ⁻⁶)	6.5e ⁻⁴ (8e ⁻⁶)
FMIN	0.44 (0.005)	0.48 (0.006)	5.6e ⁻⁴ (5e ⁻⁶)	5.3e ⁻⁴ (7e ⁻⁶)	7.5e ⁻⁴ (6e ⁻⁶)	7.2e ⁻⁴ (8e ⁻⁶)
CING-HIP L	0.30 (0.009)	0.34 (0.008)	6.9e ⁻⁴ (1e ⁻⁵)	6.5e ⁻⁴ (1e ⁻⁵)	8.3e ⁻⁴ (1e ⁻⁵)	7.8e ⁻⁴ (1e ⁻⁵)
CING-HIP R	0.29 (0.004)	0.32 (0.005)	6.8e ⁻⁴ (1e ⁻⁵)	6.3e ⁻⁴ (9e ⁻⁶)	8.1e ⁻⁴ (1e ⁻⁵)	7.6e ⁻⁴ (1e ⁻⁵)
<i>Motor</i>						
CST L	0.52 (0.006)	0.55 (0.006)	5.0e ⁻⁴ (1e ⁻⁵)	4.7e ⁻⁴ (1e ⁻⁵)	7.1e ⁻⁴ (9e ⁻⁶)	6.8e ⁻⁴ (9e ⁻⁶)
CST R	0.54 (0.004)	0.57 (0.006)	4.7e ⁻⁴ (8e ⁻⁶)	4.4e ⁻⁴ (8e ⁻⁶)	7.1e ⁻⁴ (8e ⁻⁶)	6.7e ⁻⁴ (9e ⁻⁶)
<i>Language</i>						
IFOF L ^b	0.44 (0.005)	0.48 (0.004)	5.8e ⁻⁴ (7e ⁻⁶)	5.4e ⁻⁴ (7e ⁻⁶)	7.6e ⁻⁴ (8e ⁻⁶)	7.3e ⁻⁴ (9e ⁻⁶)
IFOF R	0.45 (0.005)	0.49 (0.006)	5.6e ⁻⁴ (6e ⁻⁶)	5.2e ⁻⁴ (7e ⁻⁶)	7.6e ⁻⁴ (7e ⁻⁶)	7.3e ⁻⁴ (7e ⁻⁶)
SLF L ^b	0.44 (0.005)	0.47 (0.005)	5.4e ⁻⁴ (5e ⁻⁶)	5.1e ⁻⁴ (6e ⁻⁶)	7.0e ⁻⁴ (5e ⁻⁶)	6.7e ⁻⁴ (6e ⁻⁶)
SLF R	0.37 (0.003)	0.41 (0.005)	5.5e ⁻⁴ (4e ⁻⁶)	5.2e ⁻⁴ (5e ⁻⁶)	7.0e ⁻⁴ (5e ⁻⁶)	6.8e ⁻⁴ (6e ⁻⁶)
SLFT L ^b	0.42 (0.005)	0.45 (0.005)	5.8e ⁻⁴ (6e ⁻⁶)	5.5e ⁻⁴ (9e ⁻⁶)	7.1e ⁻⁴ (7e ⁻⁶)	6.8e ⁻⁴ (9e ⁻⁶)
SLFT R	0.41 (0.005)	0.44 (0.005)	5.8e ⁻⁴ (7e ⁻⁶)	5.4e ⁻⁴ (7e ⁻⁶)	7.3e ⁻⁴ (6e ⁻⁶)	7.1e ⁻⁴ (8e ⁻⁶)
<i>Social/Emotion</i>						
UF L	0.34 (0.005)	0.38 (0.005)	6.1e ⁻⁴ (7e ⁻⁶)	5.7e ⁻⁴ (9e ⁻⁶)	7.6e ⁻⁴ (8e ⁻⁶)	7.3e ⁻⁴ (1e ⁻⁵)
UF R	0.30 (0.008)	0.33 (0.007)	6.3e ⁻⁴ (1e ⁻⁵)	5.8e ⁻⁴ (8e ⁻⁶)	8.2e ⁻⁴ (2e ⁻⁵)	7.6e ⁻⁴ (1e ⁻⁵)
CING L	0.34 (0.007)	0.37 (0.011)	5.5e ⁻⁴ (7e ⁻⁶)	5.2e ⁻⁴ (1e ⁻⁵)	7.3e ⁻⁴ (8e ⁻⁶)	6.9e ⁻⁴ (1e ⁻⁵)
CING R	0.50 (0.007)	0.52 (0.008)	5.4e ⁻⁴ (6e ⁻⁶)	5.1e ⁻⁴ (9e ⁻⁶)	7.2e ⁻⁴ (7e ⁻⁶)	6.9e ⁻⁴ (1e ⁻⁵)
<i>Other</i>						
ILF L	0.42 (0.005)	0.45 (0.005)	6.1e ⁻⁴ (7e ⁻⁶)	5.8e ⁻⁴ (8e ⁻⁶)	7.9e ⁻⁴ (6e ⁻⁶)	7.5e ⁻⁴ (8e ⁻⁶)
ILF R	0.36 (0.003)	0.39 (0.004)	6.3e ⁻⁴ (7e ⁻⁶)	5.9e ⁻⁴ (8e ⁻⁶)	7.8e ⁻⁴ (7e ⁻⁶)	7.4e ⁻⁴ (7e ⁻⁶)
FMAJ	0.38 (0.003)	0.42 (0.005)	5.1e ⁻⁴ (7e ⁻⁶)	4.7e ⁻⁴ (6e ⁻⁶)	7.5e ⁻⁴ (6e ⁻⁶)	7.2e ⁻⁴ (7e ⁻⁶)
MID CC			4.0e ⁻⁴ (6e ⁻⁶)	3.7e ⁻⁴ (8e ⁻⁶)	7.6e ⁻⁴ (1e ⁻⁵)	7.2e ⁻⁴ (1e ⁻⁵)
<i>Cerebellum</i>						
CRBL L	0.28 (0.004)	0.31 (0.004)	5.9e ⁻⁴ (9e ⁻⁶)	5.5e ⁻⁴ (9e ⁻⁶)	7.1e ⁻⁴ (1e ⁻⁵)	6.6e ⁻⁴ (1e ⁻⁵)
CRBL R	0.26 (0.004)	0.28 (0.003)	6.1e ⁻⁴ (9e ⁻⁶)	5.7e ⁻⁴ (8e ⁻⁶)	7.2e ⁻⁴ (1e ⁻⁵)	6.8e ⁻⁴ (1e ⁻⁵)

ATR: anterior thalamic radiations; FMIN: forceps minor; CING-HIP: cingulum-hippocampus; CST: cortico-spinal tract; IFOF: inferior fronto-occipital fasciculus; SLF: superior longitudinal fasciculus; SLFT: superior longitudinal fasciculus temporal; UF: uncinate fasciculus; CING: cingulate; ILF: inferior longitudinal fasciculus; FMAJ: forceps major; MID CC: middle corpus callosum; CRBL: cerebellum.

^a Common functional association.

^b Associated with Broca's area.

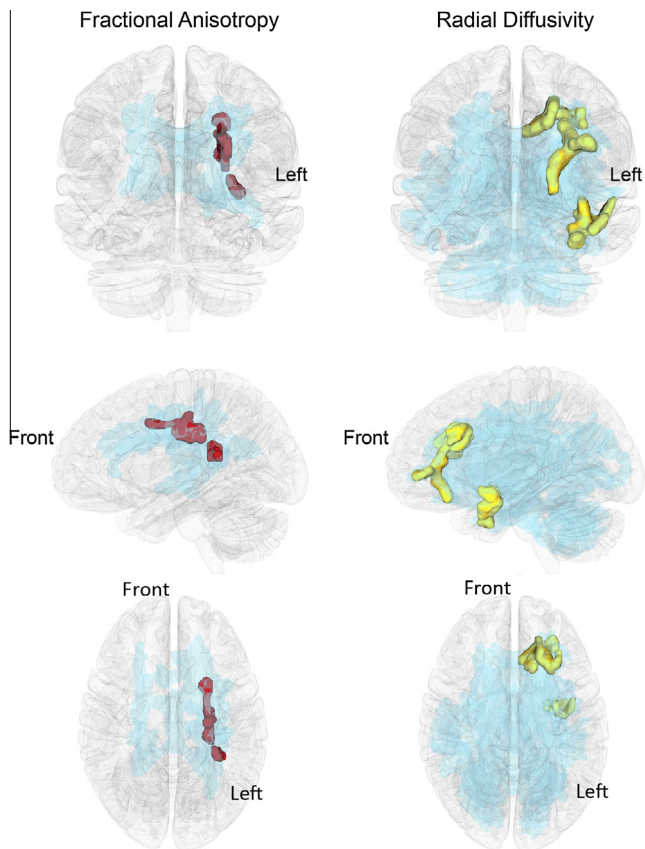


Fig. 1. Regions of significant correlations between FA (left) and RD (right) with logarithm of years in the US (blue); FA and logarithm of years in the US adjusted for reported experience in speaking English (red); RD and logarithm of years in the US adjusted for reported experience in listening to English (yellow); all thresholded at $p < 0.03$. Top row: coronal view; Middle row: sagittal view; Bottom row: axial view. See supplementary materials Fig. S1 for same data thresholded at $p < 0.05$.

(IExpAud), significant correlations ($p < 0.05$) between speaking experience and listening experience and FA were found to be located in numerous tracts in the left hemisphere as well as in the corpus callosum (Fig. S1 of the supplementary materials, red regions). Significant correlations with RD (Fig. S1 of the supplementary materials, yellow regions) and MD were more diffuse, and located in both hemispheres. No significant correlations between IExpSp or IExpAud were found with FA in the cerebellum. Additionally, no significant correlations were found between self-reported speaking or listening proficiency levels and DTI measures. See Table S5 for significant Pearson correlation coefficients for language experience and brain structural measures ($p < 0.05$, corrected for multiple comparisons and gender).

Inspection of statistical images for this correlation analysis (Fig. S1) revealed an interesting pattern of areas of increased significance, with p -values < 0.03 (shown in red and yellow in Fig. 1). These more robust correlations were found only in the left cerebral hemisphere. At this significance threshold, correlations with FA were found for speaking, but not for listening experience; whereas correlations with RD and MD were found for listening experience, but not for speaking experience. Moreover, the locations of the significant correlations with speaking and listening at a threshold of $p < 0.03$ were spatially distinct. Significant positive correlations between FA and IExpSp were found in clusters of voxels located in left posterior white matter tract regions of the left IFOF, SLF, ILF, and CST. Significant negative correlations between RD and MD and IExpAud were found in adjacent and/or overlapping clus-

Table 3

Significant correlations ($p < 0.05$) between white matter measures and log (years in the US) by tract.

Tract (by associated function) ^a	Fractional Anisotropy	Radial Diffusivity	Mean Diffusivity	Axial Diffusivity
<i>Executive</i>				
ATR L	0.589	-0.799	-0.786	-0.783
ATR R	0.376	-0.742	-0.766	-0.756
FMIN	0.615	-0.718	-0.813	-0.874
CING-HIP L	0.570	-0.750	-0.760	-0.608
CING-HIP R	0.697	-0.718	-0.755	-0.795
<i>Motor</i>				
CST L	0.784	-0.666	-0.622	-0.673
CST R	0.631	-0.671	-0.667	-0.701
<i>Language</i>				
IFOF L ^b	0.796	-0.797	-0.739	-0.698
IFOF R	0.789	-0.822	-0.816	-0.831
SLF L ^b	0.738	-0.767	-0.734	-0.699
SLF R	0.694	-0.799	-0.714	-0.690
SLFT L ^b	0.443	-0.695	-0.672	-0.575
SLFT R	0.496	-0.701	-0.686	-0.502
<i>Social/Emotion</i>				
UF L	0.539	-0.752	-0.723	-0.760
UF R	0.482	-0.670	-0.727	-0.671
CING L	0.679	-0.742	-0.674	-0.637
CING R	0.636	-0.712	-0.571	-0.571
<i>Other</i>				
ILF L	0.770	-0.770	-0.749	-0.833
ILF R	0.811	-0.791	-0.774	-0.821
FMAJ	0.848	-0.822	-0.728	-0.806
MID CC	0.730	-0.653	-0.615	-0.630
<i>Cerebellum</i>				
CRBL L		-0.702	-0.689	-0.731
CRBL R		-0.695	-0.694	-0.775

ATR: anterior thalamic radiations; FMIN: forceps minor; CING-HIP: cingulum-hippocampus; CST: cortico-spinal tract; IFOF: inferior fronto-occipital fasciculus; SLF: superior longitudinal fasciculus; SLFT: superior longitudinal fasciculus temporal; UF: uncinate fasciculus; CING: cingulate; ILF: inferior longitudinal fasciculus; FMAJ: forceps major; MID CC: middle corpus callosum; CRBL: cerebellum.

^a Common functional association.

^b Associated with Broca's area.

ters of voxels located in left anterior white matter tract regions in the IFOF, UF, ILF and adjacent regions. Table 4 provides Pearson correlation coefficients for these more robust correlation findings ($p < 0.03$, corrected for multiple comparisons and gender). Detailed MNI coordinate and voxel count information is provided in Table S4 of the supplementary materials. Scatter plots of FA and RD correlations with IExpSp and IExpAud for the left IFOF are shown in Supplementary Fig. S3.

While Tables 2–4 present results in terms of each white matter tract as a whole, it is important to note that our significant findings were localized in discrete and specific locations within these tracts. For example, Table 4 shows multiple significant correlations in the left IFOF, but the locations of significant effects are more focal. At a significance threshold of $p = 0.03$, our measure of English listening experience was significantly negatively correlated with RD and MD in the left anterior IFOF, whereas our measure of English speaking experience was positively correlated with FA in left posterior regions of the same tract. These differences in locations of correlations at a significance level of 0.03 for the different language measures and modes are illustrated in Fig. 2.

4. Discussion

In this study of bilingual adults immersed in a naturalistic L2 learning environment, we found widespread bilateral differences in DTI measures of cerebral white matter structure between Spanish-English bilingual and native English-speaking monolin-

Table 4
Significant correlations ($p < 0.03$) between white matter measures and log (years in the US) adjusted for language experience by tract.

Tract (by associated function) ^a	Speaking	Listening	
	Fractional Anisotropy	Radial Diffusivity	Mean Diffusivity
<i>Executive</i>			
ATR L		–0.599	–0.688
ATR R			
FMIN		–0.698	–0.746
CING-HIP L		–0.637	
CING-HIP R			
<i>Motor</i>			
CST L	0.703		
CST R			
<i>Language</i>			
IFOF L ^b	0.701	–0.832	–0.724
IFOF R			
SLF L ^b	0.818	–0.760	–0.704
SLF R			
SLFT L ^b			
SLFT R			
<i>Social/Emotion</i>			
UF L		–0.791	–0.758
UF R			
CING L		–0.666	–0.711
CING R			
<i>Other</i>			
ILF L	0.736	–0.746	–0.795
ILF R			
FMAJ			
MID CC		–0.547	
<i>Cerebellum</i>			
CRBL L			
CRBL R			

ATR: anterior thalamic radiations; FMIN: forceps minor; CING-HIP: cingulum-hippocampus; CST: cortico-spinal tract; IFOF: inferior fronto-occipital fasciculus; SLF: superior longitudinal fasciculus; SLFT: superior longitudinal fasciculus temporal; UF: uncinate fasciculus; CING: cingulate; ILF: inferior longitudinal fasciculus; FMAJ: forceps major; MID CC: middle corpus callosum; CRBL: cerebellum.

^a Common functional association.

^b Associated with Broca's area.

qual adults. These differences were in tracts associated with a diverse range of functions, including language processing and production, executive function, motor function, social cognition, and emotion. In the bilingual group, we also found correlations between white matter structural characteristics and duration of residence in the US, with longer residence associated with regional increases in FA, and decreases in AD, RD and MD. When duration of residence in the US was adjusted for experience in listening to and speaking English, the regions showing the most significant white matter structure correlations ($p < 0.03$) were more focal, restricted to the left hemisphere, and associated with both different regions and different measures. Increased experience in speaking English was most robustly associated with increases in FA in more posterior regions of the brain, including posterior regions of the left SLF and the left IFOF. Increased experience in listening to English was most robustly correlated with decreases in RD and MD in anterior white regions, including regions of the left uncinate fasciculus and the left IFOF.

These findings indicate that foreign language listening experience most robustly affects white matter regions associated with language motor production, and that foreign language speaking experience most robustly affects white matter regions associated with language comprehension. The results from the present study suggest that language learning induces white matter changes that enhance sensory-motor connections, with auditory English exposure more robustly affecting motor pathways, and experience in

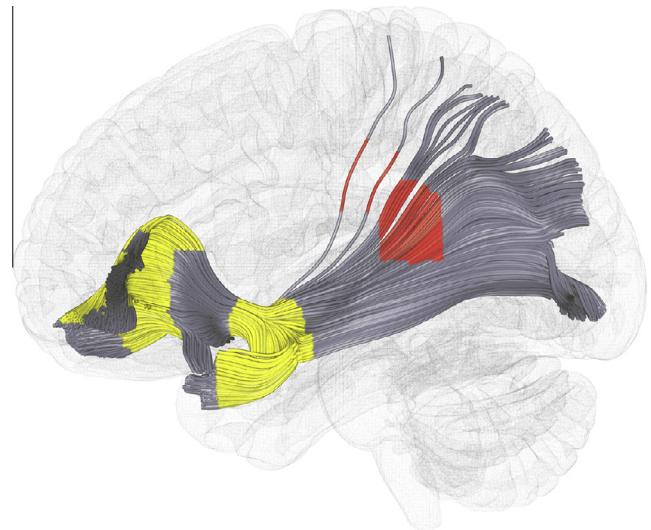


Fig. 2. View of the left IFOF showing regions of significant correlations ($p < 0.03$) with English speaking experience (red) and English listening experience (yellow). RD (yellow) and MD (not shown) are negatively correlated with English listening experience in the anterior IFOF, whereas FA (red) is positively correlated with English speaking experience in more posterior regions of this tract.

English speaking more robustly affecting pathways involved in speech perception. These findings, which at first may seem counterintuitive, parallel theories of language learning in both adults and children, including recent findings that for young children learning their first language, listening to speech activates regions involved in motor production (Kuhl et al., 2014). The findings also suggest that white matter structural features play an important role in reported behavioral and functional relationships between speech production and speech perception.

Numerous DTI studies comparing brain structure of monolinguals and bilinguals have reported significant group differences, but the direction of these differences has not been consistent. This is especially evident for white matter FA, the most commonly reported DTI measure. Cummine and Boliek, who studied a population of late sequential adult bilinguals (Cummine & Boliek, 2012), and Gold and colleagues, who studied aging sequential bilinguals (Gold et al., 2013), reported higher FA in bilinguals than in monolinguals, which is in congruence with the findings from the present study. Other studies, including one of late sequential adult bilinguals (Pliatsikas et al., 2015), and multiple studies of simultaneous bilinguals (Luk et al., 2011; Mohades et al., 2015; Pliatsikas et al., 2015), have reported findings in the opposite direction; that is, increased FA in bilinguals as compared to monolinguals. Explanations for this lack of congruence have sometimes been provided. For example, Cummine and Boliek suggest that the direction of their findings may be attributable to the relatively short time of L2 learning in their bilingual sample. This explanation may apply to our study as well, considering that many individuals in our bilingual sample had a relatively short immersion time. Gold and colleagues suggest that the direction of their findings may have been related to increased genetic risk for dementia in their bilingual sample. We propose that a broader range of environmental factors may affect baseline differences in white matter structure across groups and that these may have a more pervasive effect on structural differences between monolinguals and bilinguals than has been previously discussed. These may include environmental factors during early development, as well as native country linguistic diversity.

Another factor that may have affected the direction of our findings as compared to multiple other studies is the relatively low

immersion time for our bilingual sample (assuming an “effective” immersion time for other studies calculated by subtracting the AoA from the mean subject age). Interestingly, an analysis of our group difference data after performing a median split on the bilingual sample, based on median length of residence in the US (lower immersed subgroup with ≤ 2 years residing in the US, and a more immersed subgroup with ≥ 4 years residing in the US) revealed that, although both subgroups demonstrated mean FA values that were below the monolingual sample across virtually all white matter regions studied, these differences were significant only for the lower immersed subgroup and not significant for the more immersed subgroup (see the Supplementary Materials for more details on the median split analysis). These findings, along with the results reported in Tables 3 and 4, suggest that for bilinguals immersed in a second-language learning environment, there is a trajectory of FA increases over time. They indicate that the increased FA associated with bilingualism may take time to develop, and the shorter immersion duration of our bilingual sample may have been insufficient to produce the group differences in FA reported by other studies of bilinguals who had longer experience with the second language. It is also notable that the age of acquisition (AoA) for our bilingual sample was relatively high compared to other studies. Even though the behavioral and neuroimaging evidence indicate that there is not a strict “critical period” for second language acquisition, functional and structural imaging studies of the brain suggest that the timing of second language experience plays an important role in the acquisition of a second language (Li, Legault, & Litcofsky, 2014). For the current study, it is possible that the second language learners with higher AoA may have missed an important “sensitive” period, which may have decreased the rate at which white matter microstructure changes over time as compared to bilinguals who acquire a second language during infancy or childhood. In short, our bilingual sample differs from those of most previous studies in that it has a relatively low immersion time and relatively high AoA. The combination of these two factors may account for the different pattern of results as compared to studies of early bilingual learners with a long duration of experience with the second language.

It is well established that environmental factors, including SES and language input, affect the rate and course of native language development (Hoff, 2003, 2006; Noble, Engelhardt, et al., 2015; Raizada & Kishiyama, 2010), as well as early brain structure in areas that support language and reading (Noble, Houston, et al., 2015; Raizada, Richards, Meltzoff, & Kuhl, 2008). There is also behavioral evidence that passive exposure to a foreign language can increase sensitivity to the structure of that language in both children (Johnson, Westrek, Nazzi, & Cutler, 2011) and adults (Orena, Theodore, & Polka, 2015); this type of exposure during childhood can improve the ability to speak that language in adulthood (Au, Knightly, Jun, & Oh, 2002). Neural mechanisms presumably underlie these behavioral findings. Linguistic factors may be especially relevant to studies of bilingualism, which often include subjects from a diverse range of countries, or participants from countries where more than one language predominates. In the present study, there were no significant differences between the mean SES of the monolingual and bilingual samples, and both samples were comprised of individuals from countries where a single language (English or Spanish, respectively), is predominantly spoken. However, the effects of baseline differences in linguistic experience between samples in studies of bilingualism should be considered when comparing findings across studies, and is an area where future investigation would be valuable.

Despite the lack of consensus in the direction of white matter structural changes in the literature associated with bilingualism (Table 1), a number of specific white matter tracts have been repeatedly identified to be associated in bilingualism across these

same studies and in the present investigation. These are: the left IFOF (Gold et al., 2013; Mohades et al., 2012, 2015; Pliatsikas et al., 2015), the left SLF (Luk et al., 2011; Pliatsikas et al., 2015), and the corpus callosum (Gold et al., 2013; Luk et al., 2011). The repeated appearance of these regions, in studies of both adults and children, indicate that they play an important role in the learning and utilization of a second language across a broad age span. Furthermore, there is substantial evidence in the literature that these particular regions play important roles in language processing. According to the dual stream model of cortical organization for language processing (Hickok & Poeppel, 2004; Rauschecker & Scott, 2009), the IFOF is a major component of a ventral stream that is involved in speech comprehension, and the SLF is a major fiber tract of a dorsal stream that is associated with sensorimotor language integration. Although DTI metrics provide information about white matter structure, due to the indirect nature of the relationship between water diffusion and cellular features, care must be taken in interpreting results in terms of specific white matter features at the biological or cellular level (Jones, Knösche, & Turner, 2013). As previously noted, DTI measures are sensitive to numerous white matter structural features, including degree of myelination, axonal caliber, fiber packing density, intra-cellular volume fraction of extra neurite glial cells and proportion of crossing fibers. Decreases in RD and MD have been associated with higher density and packing of fibers, and our findings of correlations with listening experience may indicate that white matter structural changes associated with this mode of language reflect increased networking in left hemisphere language motor planning areas. FA has been associated with processing efficiency (Gold, Powell, Xuan, Jiang, & Hardy, 2007), and our findings of increased FA with increased speaking experience may indicate that speaking in the immersive environment is associated with improved processing efficiency in speech perception areas in adult bilinguals.

One limitation of this study is that residence in a foreign country is an indirect measure of foreign language immersion, and factors associated with moving to a new cultural environment unrelated to language may also impact white matter structure. Notably, as evident from inspection of Table S1, one bilingual subject who had been residing in the US for 10 years reported very low levels of English comprehension and speaking proficiency. This lack of correspondence between duration of residence and self-reports of language proficiency lends some uncertainty to the interpretation of the correlations between white matter structural measures and duration of residence alone in the bilingual sample. However, the focal nature and localization of clusters exhibiting correlations between white matter structure and speaking and listening experience to spatially distinct left hemisphere regions that have been associated with language processing and production supports the interpretation that these specific findings are attributable to language immersion.

Another limitation is the subjective nature of the language experience and proficiency measures for the bilingual sample, which were retrospective and based on self-reports. This study represents an initial step in the investigation of relationships between experience in the two language modes and white matter structure in bilinguals. Future studies would benefit from objective assessments of language proficiency. Additionally, while self-reported measures of proficiency have previously been utilized in the investigation of bilingual learning (García-Pentón, Fernández, Iturria-Medina, Gillon-Dowens, & Carreiras, 2014; Klein, Mok, Chen, & Watkins, 2014; Li, Sepanski, & Zhao, 2006; Li, Zhang, Tsai, & Puls, 2014; Marian, Blumenfeld, & Kaushanskaya, 2007; Yow & Li, 2015; Zinszer, Malt, Ameal, & Li, 2014), future work employing automated recording technology that captures detailed first-person information about the L2 language environment and speech production patterns in real-time during naturalistic immer-

sion, similar to that which has been employed in studies of language development in children (García-Sierra et al., 2011; Oller et al., 2010), could allow for the effects of these variables on white matter structural changes to be parsed out in more detail.

An important consideration in interpreting DTI measures of white matter microstructural differences between subject populations is that the diffusion signal is affected by many factors, including the presence of crossing fibers from different tracts, the fiber volume fraction, as well as the proportion of non-neuronal cells (such as glial cells) located within each voxel. Although mode of anisotropy (MO) did not reach significance for most regions of significant FA, we found that MO in clusters exhibiting significantly decreased FA was more negative in the bilingual sample than in the monolingual sample. A negative mode of anisotropy, in conjunction with decreased FA, is generally associated with a higher proportion of crossing fibers. This suggests a possibility of a difference in crossing fiber distributions between the two sample populations, but requires further investigation utilizing techniques that can more directly characterize these features.

In summary, we report widespread differences in white matter structure between adult bilinguals and monolinguals in regions of the brain associated with a diverse range of functions. In bilingual adults, our findings suggest that the amount of experience in a naturalistic L2 immersive learning environment affects the degree of white matter structural changes, where listening and speaking experience have more robust effects on white matter structure in regions of the brain associated with speech production and speech perception, respectively.

Conflict of interest

The authors declare no competing financial interests.

Acknowledgments

This study was supported by a National Science Foundation Science of Learning Center grant to the UW LIFE Center (P.K.K, PI: Grant No. SMA-0835854).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandl.2016.07.004>.

References

- Andersson, J., Jenkinson, M., & Smith, S. (2007a). FMRIB technical report: TR07JA1: Non-linear optimisation.
- Andersson, J., Jenkinson, M., & Smith, S. (2007b). FMRIB technical report: TR07JA2: Non-linear registration.
- Assaf, Y., & Pasternak, O. (2008). Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review. *Journal of Molecular Neuroscience*, 34(1), 51–61.
- Au, T. K.-F., Knightly, L. M., Jun, S.-A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science*, 13(3), 238–243.
- Callan, D. E., Jones, J. A., Callan, A. M., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage*, 22(3), 1182–1194.
- Callan, D. E., Tajima, K., Callan, A. M., Kubo, R., Masaki, S., & Akahane-Yamada, R. (2003). Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *NeuroImage*, 19(1), 113–124.
- Cummine, J., & Boliek, C. A. (2012). Understanding white matter integrity stability for bilinguals on language status and reading performance. *Brain Structure and Function*, 218(2), 595–601.
- García-Pentón, L., Fernández, A. P., Iturria-Medina, Y., Gillon-Dowens, M., & Carreiras, M. (2014). Anatomical connectivity changes in the bilingual brain. *NeuroImage*, 84(C), 495–504.
- García-Pentón, L., García, Y. F., Costello, B., Duñabeitia, J. A., & Carreiras, M. (2016). The neuroanatomy of bilingualism: How to turn a hazy view into a the full picture. *Language, Cognition and Neuroscience*, 31(3), 303–327.
- García-Sierra, A. (2007). *Language contexts in speech perception: Testing the double phonetic standard in bilinguals* Unpublished Doctoral Dissertation. Austin, TX: University of Texas.
- García-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication*, 51(4), 369–378.
- García-Sierra, A., Ramírez-Esparza, N., Silva-Pereyra, J., Siard, J., & Champlin, C. A. (2012). Assessing the double phonemic representation in bilingual speakers of Spanish and English: An electrophysiological study. *Brain and Language*, 121(3), 194–205.
- García-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., et al. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, 39(4), 546–557.
- Gold, B. T., Johnson, N. F., & Powell, D. K. (2013). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia*, 51(13), 2841–2846.
- Gold, B. T., Powell, D. K., Xuan, L., Jiang, Y., & Hardy, P. A. (2007). Speed of lexical decision correlates with diffusion anisotropy in left parietal and frontal white matter: Evidence from diffusion tensor imaging. *Neuropsychologia*, 45(11), 2439–2446.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1–2), 67–99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74(5), 1368–1378.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26(1), 55–88.
- Hollingshead, A. (2011). Four factor index of social status. *Yale Journal of Sociology*, 8, 21–52.
- Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D. S., et al. (2008). Tract probability maps in stereotaxic spaces: analyses of white matter anatomy and tract-specific quantification. *NeuroImage*, 39(1), 336–347.
- Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science*, 14(5), 1002–1011.
- Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, 73(C), 239–254.
- Klein, D., Mok, K., Chen, J.-K., & Watkins, K. E. (2014). Age of language learning shapes brain structure: A cortical thickness study of bilingual and monolingual individuals. *Brain and Language*, 131, 20–24.
- Kuhl, P. K., & Damasio, A. R. (2013). Language. In E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. Siegelbaum, & A. Hudspeth (Eds.), *Principles of Neural Science* (pp. 1353–1372). New York: McGraw-Hill.
- Kuhl, P. K., Ramirez, R. R., Bosseler, A., Lin, J. F. L., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences*, 111(31), 11238–11245.
- Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, 58, 301–324.
- Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, 38(2), 202–210.
- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(3), 673–680.
- Luk, G., Bialystok, E., Craik, F. I. M., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *Journal of Neuroscience*, 31(46), 16808–16813.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967.
- Mohades, S. G., Struys, E., Van Schuerbeek, P., Mondt, K., Van De Craen, P., & Luypaert, R. (2012). DTI reveals structural differences in white matter tracts between bilingual and monolingual children. *Brain Research*, 1435(30), 72–80.
- Mohades, S. G., Van Schuerbeek, P., Rosseel, Y., Van De Craen, P., Luypaert, R., & Baeken, C. (2015). White-matter development is different in bilingual and monolingual children: a longitudinal DTI study. *PLoS ONE*, 10(2), e0117968.
- Mori, S., Wakana, S., Nagae-Poetscher, L., & van Zijl, P. (2005). *MRI atlas of human white matter*. Amsterdam: Elsevier.
- Noble, K. G., Engelhardt, L. E., Brito, N. H., Mack, L. J., Nail, E. J., Angal, J., et al. (2015). Socioeconomic disparities in neurocognitive development in the first two years of life. *Developmental Psychobiology*, 57(5), 535–551.
- Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., et al. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, 18(5), 773–778.
- Oishi, K., Faria, A., & Mori, S. (2010, May). *JHU-MNI-ss Atlas*. <<https://http://www.slicer.org/publications/item/view/1883>>.
- Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., et al. (2010). Automated vocal analysis of naturalistic recordings from children with autism,

- language delay, and typical development. *Proceedings of the National Academy of Sciences*, 107(30), 13354–13359.
- Orena, A. J., Theodore, R. M., & Polka, L. (2015). Language exposure facilitates talker learning prior to language comprehension, even in adults. *Cognition*, 143(C), 36–40.
- Pliatsikas, C., Moschopoulou, E., & Saddy, J. D. (2015). The effects of bilingualism on the white matter structure of the brain. *Proceedings of the National Academy of Sciences*, 112(5), 1334–1337. Fig. 1.
- Raizada, R. D. S., & Kishiyama, M. M. (2010). Effects of socioeconomic status on brain development, and how cognitive neuroscience may contribute to levelling the playing field. *Frontiers in Human Neuroscience*, 4, 1–11.
- Raizada, R. D. S., Richards, T. L., Meltzoff, A., & Kuhl, P. K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *NeuroImage*, 40(3), 1392–1401.
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718–724.
- Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., & Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8), 712–721.
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., et al. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, 105(46), 18035–18040.
- Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, 17(10), 2387–2399.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155.
- Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., et al. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, 31(4), 1487–1505.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23(Suppl. 1), S208–219.
- Stein, M., Winkler, C., Kaiser, A., & Dierks, T. (2014). Structural brain changes related to bilingualism: Does immersion make a difference? *Frontiers in Psychology*, 5, 1116.
- van den Bosch, J. J. F., & Tavabi, K. (2016). *Niprov: July 2016: Versions, copies and snapshots*. Zenodo. <http://dx.doi.org/10.5281/zenodo.57177>.
- Wakana, S., Caprihan, A., Panzenboeck, M. M., Fallon, J. H., Perry, M., Gollub, R. L., et al. (2007). Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*, 36(3), 630–644.
- Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, 92, 381–397.
- Yow, W. Q., & Li, X. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: Why variations in bilingual experiences matter. *Frontiers in Psychology*, 6, 164.
- Zinszer, B. D., Malt, B. C., Ameel, E., & Li, P. (2014). Native-likeness in second language lexical categorization reflects individual language history and linguistic community norms. *Frontiers in Psychology*, 5, 1203.