Journal of Neurolinguistics 56 (2020) 100937

Contents lists available at ScienceDirect

Journal of Neurolinguistics

journal homepage: www.elsevier.com/locate/jneuroling

Onset age of second language acquisition and fractional anisotropy variation in multilingual young adults

Gigi Luk [a,*], Laura Mesite [b], Sibylla Leon Guerrero [c]

[a] McGill University, Canada
[b] MGH Institute of Health Professions, USA
[c] Harvard Graduate School of Education, USA

ARTICLE INFO

Keywords:
Bilingualism
Multilingualism
DTI
FA
TBSS
Adults

ABSTRACT

The experience of speaking two languages, particularly the age of second language acquisition (L2AoA), has been robustly associated with differences in metrics indicative of white matter structure. However, bilingual language experience is complex, described not only by acquisition but also functional language usage and proficiency. We examined whether the onset age of the *second most proficient language* (pL2AoA) would indeed be a sensitive measure that incorporates onset age of second language acquisition and self-reported L2 proficiency and correlates with variation in fractional anisotropy in white matter. Thirty-one multilingual young adults with a range of L2AoA and pL2AoA from an English-dominant community participated. After accounting for age, gender, nonverbal IQ, and parental education, both L2AoA and pL2AoA were significantly correlated with fractional anisotropy in the corpus callosum. Furthermore, we show that a larger and more extensive cluster in the corpus callosum was correlated with pL2AoA as compared to the conventional L2AoA.

1. Introduction

Brain plasticity has been associated with varied life experiences, such as stress (see McEwen, Nasca, & Gray, 2016 for a review), video-gaming experience (e.g., Kühn, Gleich, Lorenz, Linderberger & Gallinat, 2014), physical exercise (see Cassilhas, Tufik, & de Mello, 2016 for a review), literacy (e.g., Boltzmann, Mohammadi, Samii, Münte & Rüsseler, 2017), and music (see Kraus & White-Schwoch, 2017 for a review). The human brain adapts to these experiences, leading to both functional and structural changes. One notable life experience, the ability to speak and understand multiple languages, or multilingualism, has been linked to systematic variation in brain function and structure (see Li, Legault, & Litcofsky, 2014; Pliatsikas & Luk, 2016 for a review). Here, we adopt a broad definition of multilingualism to describe those who report having the ability to converse in and use multiple languages.

To date, most research on plasticity associated with multilingualism has focused on group comparisons between monolinguals and bilinguals, with a few exceptions that consider bilingualism as a continuum or a spectrum (e.g., DeLuca, Rothman, Bialystok, & Pliatsikas, 2019). However, multilingualism is complex, dynamic, and multifaceted (Grosjean, 2013), encompassing functional usage of both languages, proficiency of (second) languages, and age of second language acquisition (Luk & Bialystok, 2013; Luk, de Sa, & Bialystok, 2011). Of these variables, age of second language acquisition (L2AoA) appears to be the most robust correlate of brain

* Corresponding author. Department of Educational and Counselling Psychology, Centre for Research on Brain, Language & Music, 3700 McTavish Street, Montréal, QC, H3Y 1A2, Canada.
E-mail address: gigi.luk@mcgill.ca (G. Luk).

https://doi.org/10.1016/j.jneuroling.2020.100937
Received 2 March 2020; Received in revised form 15 May 2020; Accepted 10 July 2020
Available online 20 August 2020

function (see Cargnelutti, Tomasino, & Fabbro, 2019 for a meta-analysis) and structure (Klein, Mok, Chen, & Watkins, 2014). L2AoA conventionally denotes the age at which an individual began acquiring his or her second language. However, language acquisition is a long and individualistic process that may span years in childhood and continue into adulthood (Kidd, Donnelly, & Christiansen, 2018). Thus, while L2AoA captures the onset age at which second language exposure begins, it may not sufficiently account for other important dimensions of bilingualism, in particular the functional usage of multiple languages and second language proficiency. In this paper, we introduce the concept of proficiency ranking for age of acquisition and explore whether the onset age of acquiring *the second most proficient language* (pL2AoA) can be considered a sensitive correlate with variation in fractional anisotropy in white matter among multilingual young adults. Since pL2AoA is anchored in an individual's current language proficiency while also providing a measure of duration, we consider pL2AoA to be a rich measure of language experience beyond simple L2AoA for multilingual adults.

1.1. White matter differences between monolinguals and multilinguals

Compared to monolinguals, multilinguals who have lifelong experience using multiple languages demonstrate differences in functional brain activation and in density or volume of brain structures (see Costa & Sebastián-Gallés, 2014; Hayakawa & Marian, 2019 for reviews). In white matter analysis, which quantifies directional water diffusion in white matter tracts,[1] group comparisons between monolingual and bilingual individuals have demonstrated robust differences in white matter when groups are reliably determined to have different language experiences. However, the precise nature of these differences is unclear. Among young adults, Pliatsikas, Moschopoulou, and Saddy (2015) showed that bilinguals with immersive second language experience had higher fractional anisotropy (FA) in corpus callosum regions extending to bilateral inferior fronto-occipital fasciculi, uncinate fasciculi, and superior longitudinal fasciculi. Conversely, Cummine and Boliek (2013) reported that relative to English monolingual young adults, a group of Chinese-English bilinguals showed lower FA in the right inferior fronto-occipital fasciculus and the superior region of the right anterior thalamic radiation. Although at first glance these results seem divergent, differences in participant characteristics may account for these differences in findings (Pliatsikas & Luk, 2016). Similarly, in the aging population, Luk, Bialystok, Craik, and Grady (2011) and Gold, Johnson, and Powell (2013) reported contrasting FA results among older bilinguals as compared to age-matched monolinguals. Luk, Bialystok et al. (2011) reported higher FA in regions similar to those reported in Pliatsikas et al. (2015) while Gold et al. (2013) reported lower FA in similar regions. Importantly, the adults recruited in Luk and colleagues' study were around 70 years old whereas the participants in Gold and colleagues' study were younger. Recently, Anderson, Grundy, et al. (2018) reported that in a slightly older sample, instead of matching on a single variable like Luk, Bialystok, et al. (2011) and Gold et al. (2013), multivariate propensity score matching showed that no group difference was observed in FA, but there were group differences in axial diffusivity in the left superior longitudinal fasciculus, with the bilingual older adults showing higher axial diffusivity values.

Relatively little work has been done on white matter comparisons in younger multilinguals and monolinguals. Mohades et al. (2012) reported cross-sectional group differences among monolingual, simultaneous bilingual, and sequential bilingual children as well as longitudinally through early adolescence (2015). Instead of whole-brain white matter, four specific language-related white matter tracts were compared: the arcuate fasciculi, superior longitudinal fasciculi, inferior fronto-occipital fasciculi and corpus callosum. While the three groups had comparable FA in the left arcuate fasciculus, left superior longitudinal fasciculus, and the anterior midbody of the corpus callosum extending to the premotor and supplementary motor cortices, simultaneous bilingual children had higher FA than the other two groups in the left inferior fronto-occipital fasciculus and lower FA than monolinguals in the anterior corpus callosum projecting to the orbital lobe. About two years later, Mohades et al. (2015) had the children return for a second scan and examined FA in the four specific tracts. In this longitudinal follow-up, simultaneous bilinguals continued to show higher FA than monolinguals in the left inferior fronto-occipital fasciculus with the sequential bilinguals not significantly different from the other groups. Interestingly, the rate of change in FA over the two years was the largest among the sequential bilinguals and the smallest in the simultaneous bilinguals.

Taken together, these group comparison studies indicate that reliable differences can be observed between monolinguals and bilinguals. However, the nuances of each study have implications for the interpretation of findings because brain outcomes reflect individual manifestations of multilingualism as an experience across the lifespan. Crucially, multilingualism is cumulative and interactional, and only prolonged multilingual experience and sustained interactions involving more than one language appear to result in structural plasticity in the brain. Further, in behavioral research comparing monolinguals and bilinguals, there have been wide variations in labeling and describing these two categories of both child and adult participants (Surrain & Luk, 2019) while only a small portion of studies describe the sociolinguistic context where multilingualism occurs. Thus, although dichotomous bilingual to monolingual comparisons provide valuable information about differences in white matter structures associated with multilingualism, examining the association between key measures of multilingual experience and white matter structure may be more meaningful.

1.2. Association between multilingual experience and white matter variation

Given the complexity of multilingual experience, studies that compare multilinguals who have different experiences seek to examine a "gradation" of such experience. Age of second language acquisition (L2AoA) has been shown to be one such gradation that allows the investigation of white matter correlates. Hämäläinen, Sairanen, Leminen, and Lehtonen (2017) compared early and late

[1] Here, we report the measurement units of white matter analysis, namely fractional anisotropy (FA), rather than using the term "integrity" to maintain transparency. This is following the cautionary note of Jones, Knösche, and Turner (2013).

Finnish-Swedish-English multilinguals in four specific white matter tracts in each hemisphere (three segments of arcuate fasciculi and inferior fronto-occipital fasciculi), all brain structures directly relevant to language processing. Although there were no significant whole-brain differences between early and late multilinguals, tract-specific results showed that early, as compared to later, multilinguals had higher FA in the frontal part of the inferior fronto-occipital fasiculus, particularly in the right hemisphere, consistent with the group comparisons findings between monolinguals and bilinguals. Kuhl et al. (2016) compared whole-brain white matter structure in young monolingual and bilingual adults. Group comparisons showed that monolinguals had higher FA than bilinguals in most white matter tracts. The bilinguals in the study were Spanish-English bilinguals who were recent immigrants. Strong positive correlations with their length of stay in the U.S. (an English-dominant country) suggested that length of immersive experience contributed to variation in FA and diffusivity of widespread white matter structures. In a similar study, Nichols and Joanisse (2016) recruited Chinese-English bilinguals who were not fully immersed in their L2 environment and showed that L2AoA positively correlated with FA in left corpus callosum, left arcuate fasiculus, and bilateral inferior longitudinal fasciculus when controlling for L2 proficiency. In a study conducted in the U.K., L2AoA was also shown to be positively correlated with FA in the corpus callosum in a group of multilinguals in England (DeLuca et al., 2019). By contrast, Rossi, Cheng, Kroll, Diaz, and Newman (2017) reported a negative correlation between L2AoA and FA in corpus callosum, bilateral inferior fronto-occipital fasciculi, and inferior longitudinal fasciculi in a group of English speakers who were learning Spanish in the U.S. This group also displayed higher FA values in several tracts when compared to their English monolingual peers. These discrepant findings in different locations, direction of relationship, and sample characteristics, bear further investigation and frame our current study.

1.3. The current study

Collectively, while it is accepted that multilingual experience correlates with variation in white matter structure, the nature and direction of the relationship are still open questions. Importantly, what is the direction of the relationship between L2AoA and FA

Table 1
Demographic and behavioral characteristics of the sample (n = 31).

	Mean	SD	Min	Max
Demographics				
Age (years)	25.16	3.76	18	33
Participant Education BSMSS Score	19.16	2.15	15	21
Parent Education BSMSS Score	17.56	4.71	3	21
English proficiency measures				
KBIT-2 Non-Verbal IQ (Standard Score)	109.71	13.22	90	130
KBIT-2 Verbal IQ (Standard Score)	109.74	18.27	59	160
Applied English skills in WMLS-R	121.59	11.21	89	145
Bilingual history and non-English language usage				
Onset age of L2 acquisition (years)	7.55	3.65	2	15
Onset age of pL2 acquisition (years)	8.81	4.49	0	18
% daily listening in a non-English language	13.70	18.09	0	52
% daily speaking in a non-English language	12.87	19.14	0	70
% daily reading in a non-English language	9.94	15.45	0	52
% daily writing in a non-English language	8.71	16.54	0	59

Contexts where L2 was acquired	**Percentage of participants selecting ratings by category**				
	None at all	A little	A moderate amount	A lot	A great deal
Classroom	0%	3.2%	9.7%	9.7%	77.4%
Family	41.9%	22.6%	12.9%	6.5%	16.1%
Friends	12.9%	25.8%	16.1%	12.9%	32.3%
Reading	9.7%	19.4%	12.9%	16.1%	9.7%
Self-study	41.9%	19.4%	12.9%	16.1%	9.7%
Television	29.0%	25.8%	0%	25.8%	19.4%
Work	51.6%	9.7%	3.2%	12.9%	22.6%

Contexts where pL2 was acquired	**Percentage of participants selecting ratings by category**				
	None at all	A little	A moderate amount	A lot	A great deal
Classroom	3.2%	0%	9.7%	6.5%	80.7%
Family	58.1%	16.1%	12.9%	6.5%	6.5%
Friends	6.5%	25.8%	22.6%	16.1%	29.0%
Reading	12.9%	12.9%	25.8%	29.0%	19.4%
Self-study	32.3%	19.4%	19.4%	19.4%	9.7%
Television	19.4%	29.0%	12.9%	22.6%	16.1%
Work	54.8%	12.9%	3.2%	12.9%	16.1%

among multilingual young adults in the U.S. context? Second, how does L2AoA reflect individuals' dynamic and complex language acquisition history and proficiency? As the most common and prevalent marker of L2 acquisition, L2AoA marks the starting point of L2 exposure and usage. Nichols and Joanisse (2016) interpreted the positive relationship between L2AoA and FA in corpus callosum, arcuate fasciculus and bilateral inferior longitudinal fasciculi as the independent effect of proficiency, while DeLuca et al. (2019) suggested that the positive relationship may have been confounded with length of immersion. Clarifying the nature of how L2AoA correlates with FA in white matter microstructure will further current knowledge on how key factors in multilingual experience relate to neural plasticity.

In the current study, we examined the correlation between L2AoA (measured in a conventional manner) in a group of multilingual young adults and FA in the whole brain. Moreover, we tested the extent to which the onset age of acquiring the *second most proficient language* (pL2AoA) may be a more comprehensive and sensitive measure incorporating functional usage, L2 proficiency, and L2AoA that reflects the complexity of multilingualism in adulthood. Generally, a positive relationship is expected between functional usage of L2 and L2 proficiency as more usage leads to higher proficiency (Luk & Bialystok, 2013). However, a negative relationship is expected between L2AoA and L2 proficiency (Luk, Bialystok, et al., 2011) as the earlier L2 is acquired, the higher L2 proficiency is attained. Given multilingual experience encompasses separable but correlated dimensions (Yow & Li, 2015), L2AoA alone may not sufficiently capture the collective influence of these dimensions as L2AoA only marks the beginning of second language acquisition without taking proficiency level into account. In order to capture these overlapping constructs, we adopted Anderson, Mak, Keyvani Chahi, & Bialystok (2018) language and social background questionnaire in which participants were asked to first rank their spoken languages in order of proficiency before reporting the age at which they began acquiring each language. By capturing AoA of the *second most proficient language*, pL2AoA accounts for proficiency, functional L2 usage, and language dominance in multilinguals, thereby enriching L2AoA to serve as a measure of multilingual experience.

We expected L2AoA to be a significant correlate of FA in white matter structure, particularly in corpus callosum and bilateral inferior fronto-occipital fasciculi, two structures that were observed in most studies comparing monolinguals to multilinguals. Furthermore, we expected pL2AoA to be a sensitive measure that could correlate with broader white matter regions as an indicator of both proficiency and functional language usage, complementing the conventional measure of L2AoA which only marks the beginning of second language acquisition. Finally, we anticipated a negative relationship such that the earlier L2AoA or pL2AoA correlated with higher FA in the white matter structures.

2. Methods

2.1. Participants

Thirty-one young adults, ages 18 to 33 (M_{age} = 25.16 years, 5 male), self-reported to have exposure to a second language, were recruited via flyers and Facebook posts to participate in a neuroimaging study on language and learning at the Harvard University Center for Brain Science (CBS). Table 1 presents the demographic characteristics of the sample. Participants' language background was assessed using the language and social background questionnaire (Anderson, Mak, et al., 2018). A majority of participants were born in the U.S. ($n = 21$); other home countries included China ($n = 4$), Thailand ($n = 2$), Canada ($n = 1$), Korea ($n = 1$), Russia ($n = 1$), and the Netherlands ($n = 1$). Those born outside of the U.S. had moved to the U.S. within the previous 9.63 years on average (SD = 12.73, range = 0–32 years). The language environment where these participants were recruited was linguistically diverse although English was used predominantly in most social and professional settings.

Approximately 71% of participants ($n = 22$) reported using English as their dominant home language for the previous five years. Other dominant home languages included Mandarin ($n = 3$), Korean ($n = 2$), Thai ($n = 2$), Bengali ($n = 1$), and Gujarati ($n = 1$). On average, participants reported that 13% of their daily speaking occurred in a non-English language, with responses ranging from 0% to 52%. Approximately 61% of participants ($n = 19$) reported that both of their parents spoke English as their strongest language, while 39% ($n = 12$) indicated that their mother's and/or their father's strongest language was not English. A majority of participants reported speaking English as their first language ($n = 20$), while other first languages included Mandarin ($n = 3$), Cantonese ($n = 2$), Thai ($n = 2$), Korean ($n = 2$), Bengali ($n = 1$), and Gujarati ($n = 1$). Participants reported speaking a variety of second languages, including English ($n = 11$), Spanish ($n = 9$), French ($n = 4$), Hebrew ($n = 2$), Dutch ($n = 1$), Italian ($n = 1$), Latin ($n = 1$), and Tigrinya ($n = 1$). Nineteen participants also reported at least basic knowledge of a third language, including Spanish ($n = 7$), French ($n = 6$), Mandarin ($n = 2$), Cantonese, ($n = 1$), Hebrew (n = 1), and Portuguese (n = 1). Lastly, seven participants also reported at least basic knowledge of a fourth language, German ($n = 1$), Latin ($n = 1$), Mandarin ($n = 1$), Spanish ($n = 1$), Thai ($n = 1$), and Yiddish ($n = 1$). Self-reported L2 proficiency ranged from beginner (19.4%), beginner/intermediate (12.9%), intermediate (12.9%), intermediate/advanced (3.2%), advanced (25.8%), advanced/native (9.7%), and native (16.1%).

Given that the age at which participants began speaking their second language is one of our key question predictors, we report both the onset ages at which the participants were (1) first exposed to their second language (L2AoA), as well as (2) first exposed to their second most proficient language (pL2AoA). To obtain pL2AoA, participants were asked to rank their spoken languages from highest to lowest in order of proficiency, as well as the age at which they began speaking each language. Overall, 10 of the 31 participants reported a different L2 according to acquisition order versus according to proficiency level. In terms of self-rated speaking proficiency, given that the study took place in the U.S., a majority of the sample ($n = 25$) indicated English as their strongest language, while others indicated Mandarin (n=3), Thai ($n = 2$), and Korean ($n = 2$).[2] Since the sample generally reported English to be their dominant and strongest language and were living in a predominantly English-speaking community, we considered the sample's overall experience with immersion to be more similar to the Rossi et al. (2017) sample than to the experience of a long-time immersed minority language sample such as that of Nichols and Joanisse (2016) or De Luca et al. (2019). Therefore, we maintained our expectation that a negative relationship should be observed between FA values and L2AoA or pL2AoA, mirroring findings from previous studies with similar characteristics (Rossi et al., 2017).

2.2. Procedure and measures

Participants completed one testing session at the Harvard CBS. During this session, they provided informed consent; were familiarized with scanning procedures through a mock scanner; took part in a fMRI session followed by cognitive, language and reading assessments; and lastly, completed the language and social background questionnaire. Instructions, administration, and all testing materials were in provided English only. The fMRI session also included functional and structural runs not reported in the current manuscript. The study was approved by the Committee on the Use of Human Subjects at Harvard University.

2.2.1. Language and social background questionnaire (adapted from Anderson, Mak, et al., 2018)

In the language and social background questionnaire, participants were asked to provide information about their demographics, language acquisition history, daily usage of all languages, and self-rated language proficiency. Using these questions, two key variables, namely the onset age of L2 acquisition (L2AoA) and the onset age of acquiring the second most proficient language (pL2AoA) were obtained. In addition, participants were asked to indicate their mother's and father's highest level of education as well as their own highest level of education. Using the Barratt Simplified Measure of Social Status (BSMSS, Barratt, 2006), we calculated a composite score for parental education that ranged from 3 (both parents completed less than seventh grade) to 21 (both parents obtained a graduate degree). The average parental education composite score in our sample was 17.56 (SD = 4.71), which is roughly equivalent to both parents receiving an undergraduate degree. Relative to their parents, our sample had a slightly higher BSMSS score in terms of education, with an average of 19.16 (SD = 2.15).

2.2.2. Kaufmann Brief Intelligence Test, 2nd Edition (KBIT-2, Kaufman & Kaufman, 2004)

We administered the Matrices subtest from the as an indicator of participants' non-verbal IQ. In this task, participants were asked to choose a picture that best fit the missing space of a series of pictures. Standard scores on this measure ranged from 90 to 130, with an average of 110 ($SD = 13.22$). In addition, we administered the Verbal Knowledge subtest and the Riddles subtest as an indicator of the participants' verbal IQ (in English). The Verbal Knowledge subtest is a picture vocabulary test in which participants are asked to identify a picture shown by the experimenter. The Riddles subtest requires participants to listen to descriptions of a target word in order to identify the target word. While the Verbal Knowledge subtest assesses receptive knowledge in English, the Riddles subtest assesses expressive verbal ability in English. Standard scores of the composite of these two subtests ranged from 59 to 160, with a mean of 110 ($SD = 18.27$). Both verbal and nonverbal subtests have a reported population mean of 100 with standard deviation of 15. The slightly higher standard deviation in the verbal composite may be due to participants' diverse language backgrounds. The sample had slightly higher average verbal and nonverbal IQ than the population mean, which is not unexpected for young adults in a college community.

2.2.3. English Applied Language Proficiency

We administered the Dictation, Understanding Directions, Story Recall, and Passage Comprehension subtests from the Woodcock-Muñoz Language Survey-Revised Normative Update (WMLS-R NU, Schrank, Wendling, & Alvarado, 2010) to calculate an Applied Language Proficiency composite score for each participant. In the Dictation subtest, participants were asked to provide written responses to questions related to spelling, punctuation, capitalization, and word usage. The Understanding Directions subtest measured broad oral language skills in English including listening comprehension, lexical knowledge, and working memory. Participants were expected to point to objects on a page following verbal prompts presented. The Story Recall subtest further required participants to retell a passage from audio recordings. The number of details recalled was recorded and scored accordingly. Finally, Passage Comprehension assessed participants' reading ability in English. Together, these four subtests provided a general assessment of the language and literacy skills in English of the sample. As indicated in the assessment manual, this composite represents "the proficiency with which an individual can effectively apply listening, speaking, reading, writing, and comprehension abilities" in English (Schrank et al., 2010, p. 14). Standard scores on this measure ranged between 89 and 145 in our sample, with an average score of 122 (SD = 11.21). Participant demographic and background information is presented in Table 1.

[2] Eleven participants reported Spanish as their second strongest language, and other reported languages included English ($n = 6$), French ($n = 6$), Latin ($n = 2$), Mandarin ($n = 2$), Cantonese ($n = 1$), Gujarati ($n = 1$), Italian ($n = 1$), and Korean ($n = 1$). Third languages in terms of strength included Spanish ($n = 7$), French ($n = 6$), Mandarin ($n =$), Cantonese, ($n = 1$), Hebrew ($n = 1$), and Portuguese ($n = 1$), and fourth languages in terms of strength included German ($n = 1$), Latin ($n = 1$), Mandarin ($n = 1$), Spanish ($n = 1$), Thai ($n = 1$), and Yiddish ($n = 1$).

2.3. DTI data acquisition and analysis

MRI data were acquired on a 3T Siemens MAGNETOM Prisma Scanner (Siemens Medical Systems, Erlangen, Germany) using a 32-channel head coil. Participants were outfitted with fMRI-Compatible Insert Earphones from Sensimetrics Corporation to enable high-quality auditory presentation during the video task while protecting participants' hearing for administering instructions and other functional tasks involving processing spoken language. A standard multi-echo, diffusion weighted echo planar imaging sequence was used with the following parameters: TR = 3700 ms, TE = 83 ms, voxel size = 1.8 × 1.8 × 1.8 mm (5.83 mm^3), FoV = 209 mm × 209 mm^2, number of slices = 81, b value = 1000 s/mm^3, A » P phase encoding direction, number of diffusion gradient directions = 67, and a multi-band acceleration factor of 3.

2.3.1. Pre-processing

DTI data were first reconstructed using dcm2niix (Li, Morgan, Ashburner, Smith, & Rorden, 2016) and then processed and analyzed using FSL (5.0.11; Smith et al., 2004). After a visual inspection to determine sufficient image quality, the raw data were corrected for the effects of head motion and eddy currents using the eddy_correct command from the FMRIB diffusion toolbox with the first volume used as the reference. Next, a brain mask was generated for each participant using the bet2 command of a B = 0 image using a fractional intensity threshold of 0.3, vertical gradient of 0, and robust brain center estimation. All brain masks were visually inspected and, if necessary, manually edited in FSLeyes to ensure whole-brain coverage. Afterwards, a diffusion tensor model was fit at each voxel for each participant using DTIFIT. Then visual inspection was performed for all participants in FSLeyes to ensure that vectors were properly oriented and there were few, if any, fractional anisotropy (FA) values greater than 1 or less than 0, which would indicate a problem with alignment or a low SNR.

2.3.2. Tract-Based Spatial Statistics

Tract-Based Spatial Statistics (TBSS; Smith et al., 2006) in FSL was used to analyze the FA data. During this process, participants' data were aligned to 1 mm^3 standard space using the built-in FMRIB58_FA template and FNIRT nonlinear registration tool (Andersson, Jenkinson, & Smith, 2007). Then an average FA image was created and thinned to represent the centers of all tracts shared among participants using the common threshold of 0.2. A mask of this skeletonized image was then generated, and each participant's aligned FA data was projected onto it to create a single file containing each subjects' FA data within these common tracts.

2.3.3. Analysis

Whole-brain analyses were carried out using the randomize command (Winkler, Ridgway, Webster, Smith, & Nichols, 2014) in FSL using Threshold Free Cluster Enhancement (TFCE), demeaned data and covariates, and 5000 permutations per comparison. A TFCE-corrected p-value of 0.05 was used to determine statistical significance. These analyses investigated the presence of an association (either positive or negative) between each of our variables of interest (L2AoA, pL2AoA) and FA, controlling for the effects of biological sex, age, parental education, and non-verbal IQ. Mean FA values in the largest clusters showing significant correlations for each participant were extracted and then used to calculate partial correlations accounting for the aforementioned nuisance variables

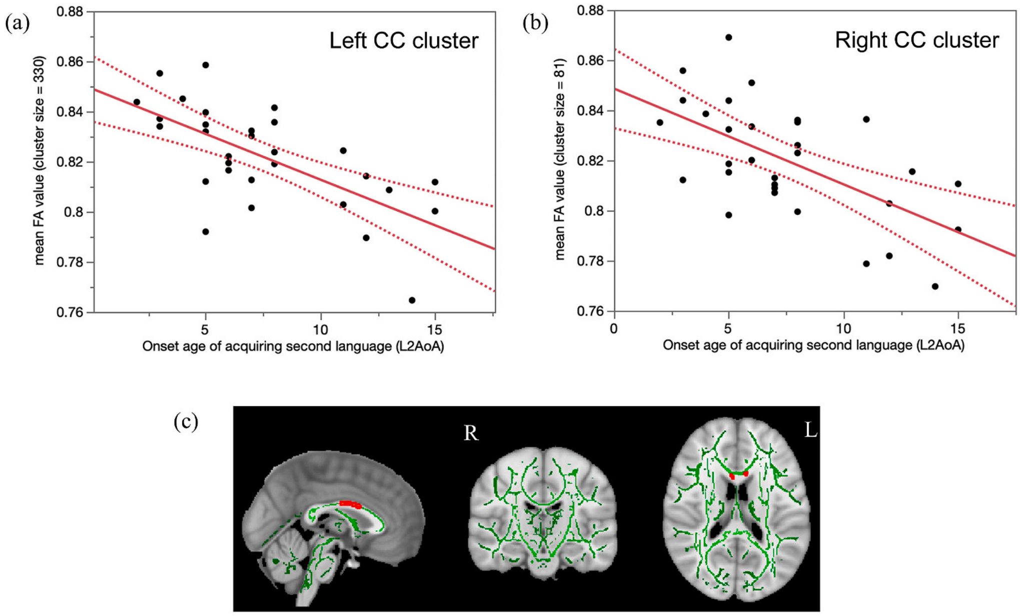

Fig. 1. Relationship between onset age of acquiring second language (L2AoA) and mean FA values in (a) cluster in left corpus callosum (MNI: 6, 5, 26, size = 330 voxels) and (b) cluster in right corpus callosum (MNI: 6, 8, 24, size = 81 voxels). The slices showing these clusters are presented in (c).

offline.

3. Results

Accounting for sex, age, parental education, and non-verbal IQ, we found a statistically significant negative association between age of second language acquisition (L2AoA) and FA in the body of the corpus callosum. The analysis revealed two significant clusters in this region, one in the left side of the corpus callosum (330 voxels, partial correlation $r = -0.74$, maximum $t = -5.22$, $p = 0.044$, MNI coordinates of cluster maximum = −6, 5, 26) and the other in the right side of the corpus callosum (81 voxels, partial correlation $r = -0.70$, maximum $t = 3.75$, $p = 0.048$, MNI coordinates of cluster maximum = 6, 8, 24). Furthermore, when including the controls described above, we also found a significant negative correlation between the onset age of acquiring the second most proficient language (pL2AoA) and FA in the body and genu of the corpus callosum extending into the corona radiata bilaterally (4404 voxels, partial correlation $r = -0.73$, maximum $t = -4.86$, $p = 0.037$, MNI coordinates of cluster maximum = −8, 12, 23) as well as the retrolenticular part of the right internal capsule (38 voxels, partial correlation $r = -0.76$, maximum $t = -5.41$, $p = 0.049$, MNI coordinates of cluster maximum = 26, −22, 8). These significant regions, enlarged for visualization using tbss_fill, are overlaid on the sample mean FA skeleton and displayed on the MNI152 template are presented in Figs. 1 and 2. There were no other significant correlations in either direction.

4. Discussion

Multilingualism is an experience-dependent process that is dynamic and complex, particularly in adulthood where language usage may be different from that in childhood. In the present study, we found that the onset age of acquiring a second language (L2AoA) was negatively correlated with FA in corpus callosum, such that higher FA was observed in those who acquired a second language earlier. Moreover, the onset age of acquiring the second most proficient language (pL2AoA) was observed to have a similar correlation, but the regions in the corpus callosum that were significantly associated with pL2AoA were more extensive than those associated with L2AoA. These findings add to the current literature on multilingualism and variation in brain structure. Importantly, L2AoA as a conventional measure documenting the onset of L2 acquisition seems to be more complex than expected. In the current study, pL2AoA accounts for both proficiency and language dominance in multilinguals and enriches L2AoA to capture functional L2 usage and perceived L2 proficiency.

A notable finding in the current study was the negative relationship between FA in corpus callosum and both L2AoA and pL2AoA. Intuitively, if multilingualism is a cumulative experience, we expect a *negative* linear relationship between L2AoA (or pL2AoA) and FA values in regions that have previously been shown to differ between monolinguals and multilinguals. Earlier L2AoA, corresponding to greater multilingual experience, is thus expected to be associated with higher FA values in these regions. Indeed, this expectation was confirmed both in our current sample and in prior studies with samples similar to ours (e.g. Rossi et al., 2017). However, our findings do not align with other studies reporting a *positive* relationship between L2AoA and FA in white matter structures (DeLuca et al., 2019; Kuhl et al., 2016; Nichols & Joanisse, 2016). We consider three explanations for the mixed results, namely the white matter tracts where correlations were observed, participant characteristics, and language dominance.

In our study, the genu and body of corpus callosum extending to bilateral corona radiata emerged as white matter tracts displaying significant negative correlations with L2AoA and pL2AoA, consistent with studies showing group differences between monolinguals and bilinguals in young (Pliatsikas et al., 2015) and older adults (Luk, Bialystok, et al., 2011). In the three studies showing a positive relationship, participants who acquired a second language *later* had higher FA values than those who acquired a second language earlier. In DeLuca et al. (2019), the callosal white matter tracts with significant L2AoA correlations were almost the same as those reported in our study. Similarly, Kuhl et al. (2016) showed that the length of stay in the U.S. (where the participants resided at the time of the study) correlated with more widespread white matter structure differences (as shown in Fig. S1 in the supplementary materials of the study), centering in the corpus callosum regions. Finally, the left corpus callosum was one tract Nichols and Joanisse (2016) reported as uniquely predicted by L2AoA (controlling for L2 proficiency).

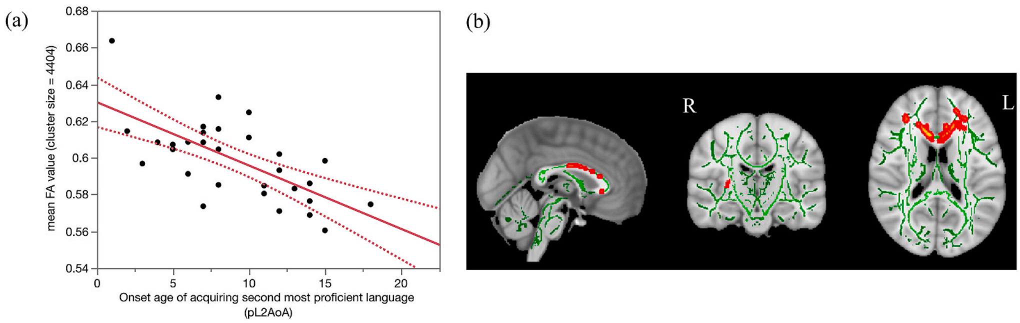

Fig. 2. (a) Relationship between onset age of acquiring second *most proficient* language (pL2AoA) and mean FA values in cluster in corpus callosum (MNI: 8, 12, 23, size = 4404 voxels). The slices showing these clusters are presented in (b).

Other white matter tracts associated with multilingual experience include the inferior longitudinal, inferior fronto-occipital, arcuate and uncinate fasciculi. Rossi et al. (2017) examined correlations between L2AoA and regions showing group differences between English monolinguals and native English speakers who learned Spanish. Negative correlations were observed in the left inferior fronto-occipital fasciculus, uncinate fasciculus, and inferior longitudinal fasciculi. In contrast, Nichols and Joanisse (2016) found positive L2AoA associations with bilateral inferior longitudinal fasciculi and left arcuate fasciculus. FA differences between simultaneous and sequential multilingual young adults (Hämäläinen et al., 2017) and children (Mohades et al., 2012) were found in right and left inferior fronto-occipital fasciculus respectively when compared to sequential multilinguals (and to monolinguals in Mohades et al., 2012). In sum, there are more commonalities than differences in terms of the tracts where the significant correlations were

Table 2
Participant characteristics and multilingual experiences examining correlations between FA in white matter tracts and L2AoA.

	Age[a] (range)	Spoken languages	Range of L2AoA[b] (range)	Correlation between FA and L2AoA	L2 proficiency	Qualitative description
DeLuca et al. (2019) N = 65	31.7±7.2 (18–52)	Various L1, English L2, some have L3	8.51±4.9 (0–22)	+	High-intermediate to high-proficiency in English based on QPT	- Mostly immigrants to England, except for 3 who were born in English-speaking countries - Self-reported usage showed mixed usage of L2 outside of home, but primarily L1 usage at home.
Kuhl et al. (2016) N = 16	26.9±3.3 (22–32)	Spanish L1, English L2	Ranges from 3 to 6 to 27-30	+	Self-report English comprehension and speaking were 4.2 and 4.3 respectively[c]. (5 = native-like proficiency)	- Immigrants to US during adulthood except for one who was born in the US - Community sample in Seattle, WA - Years residing in US ranged from 0.1 to 22 years[c] - Reported 59% (SD = 26%) and 54% (SD = 24%) listening and speaking English on average[c]
Nichols and Joanisse (2016) N = 22	22.2±4.2 (18–35)	Mandarin L1, English L2	13.8±7.1 (4–30)	+	68.3±6.6 (range: 54.2–77.1) percentage based on the Test of English as a Second Language from ETS	- Community sample in London, Ontario, Canada - Authors considered participants were "not fully immersed in their L2 from the time of first exposure" - no reported quantity or context of daily L2 usage
Rossi et al. (2017) N = 25	18–27	English L1, Spanish L2	12.1±3.5 (2–19)	–	0.55±0.22 (range: 0.2–1.0) based on proportion of accuracy in a picture naming task in Spanish. Self-rating of Spanish is 7.2±1.2 (range: 5–10 out of 10)	- Community sample in State College, PA - Native speakers of English learning Spanish - Dominant English speakers - no reported quantity or context of daily L2 usage
Current study N = 31	25.2±3.8 (18–33)	Mostly English L1, various L2 (majority were Spanish)	7.6±3.7 (2–15)	–	Beginner (19.35%) Beginner/Intermediate (12.90%) Intermediate (12.90%) Intermediate/Advanced (3.23%) Advanced (25.81%) Advanced/Native (9.68%) Native (16.13%)	- Community sample in Boston, MA. Mostly born in the U.S. - primarily native speakers of English - wide range of daily L2 usage (0–70% in speaking a non-English language)

[a] Mean age in years was reported based on the information in the publication.
[b] L2AoA was reported in terms of years based on the information in the publication.
[c] Data were calculated from the supplementary materials associated with the publication.

observed, despite the fact that the directions of the relationships are not always the same.

Next, we consider the participants' characteristics and their multilingual experiences across these studies. Table 2 presents the five studies showing divergent relationships between FA and multilingual experience. First, the three studies showing a positive relationship included multilinguals who had extensive L2 immersive experience. In contrast, the sample in the current study is more similar to Rossi et al.'s (2017) sample where the participants resided in communities where their L1 was spoken. Although it was possible that they may have had prior L2 immersive experience, this was not the case at least at the time when the study was conducted. Therefore, we speculate that it may be the nature, duration, and timing of L2 immersive experience that are key factors driving the different correlations.

One model which may synthesize the seemingly mixed findings in the literature is Pliatsikas' (2020) Dynamic Restructuring Model (DRM) which addresses the dynamic experience of second language *immersion* and its relationship to brain structure including white matter tracts, cortical gray matter, and subcortical gray matter. The DRM's Stage 1 marks initial exposure to L2, which is proposed to profoundly affect cortical and subcortical gray matter but not white matter. In Stage 2, the L2 immersive experience is consolidated, impacting white matter microstructure with specific subcortical gray matter modulations. Stage 3 denotes the peak efficiency phase where L2 immersion sustains prolonged experience. The DRM thus characterizes multilinguals' immersive experience and allows for predictions about variation in brain structure and the quality of multilingual experience.

How the active ingredient of L2 immersion results in such drastic differences in neural plasticity may be related to contextual demands for using the L2. When immersed in a L2 environment, an individual has a constant need to be vigilant about language use, particularly focusing on their less fluent second language. The positive correlations between FA in various white matter tracts and L2AoA observed in DeLuca et al. (2019), Kuhl et al. (2016), and Nichols and Joanisse (2016) may reflect the intensity of prolonged L2 immersive experience in DRM's Stage 3. However, the demands of L2 use in a L1-dominant context, as in Rossi et al. (2017) and the current study, suggest a simpler, more intuitive relationship between brain structure and multilingual experience. Turning to the literature on second language learning in a classroom context, Schlegel, Rudelson, and Tse (2012) collected monthly DTI scans on English-speaking college students enrolled in an intensive Chinese class across a period of nine months (meeting nine times for 7.5 h per week). Compared to the age- and gender-matched controls who did not enroll in any foreign language class, students of Chinese had a significant increase in FA in the white matter tracts connecting the corpus callosum extending to the temporal region, with the results being most pronounced in the genu of the corpus callosum. These findings were correlated with the course grade in the intensive Chinese class achieved by the participants and mirrored the locations of higher FA in the current study, particularly for the correlations with pL2AoA (See Fig. 1c and d) suggesting that increased proficiency in a second language played a role in white matter plasticity. Indeed, the sample in Schlegel et al. (2012) did not have immersive experience as the L2 experience was only confined to the classroom. Similarly, a majority of our sample reported learning L2 or pL2 in a classroom context (77% and 81% respectively), highlighting the possible role of L2 learning context.

Without experiencing L2 immersion in a sustained manner, it is likely that the majority of our sample of bilinguals had more restricted use of L2 in their daily life. In particular, participants' language exposure in the daily environment was not the same as those in other studies who experienced L2 immersion. The divergent results we describe prompt a more comprehensive report of participant characteristics as well as the investigation of the diverse qualities of multilingual experience in adulthood. In their review of research on the cognitive consequences of multilingualism, Fricke, Zirnstein, Navarro-Torres, and Kroll (2019) highlighted the complexity of multilingual experience and called for the investigation of cognitive correlates to be considered in light of the wide variations that exist across language processes and contexts. A similar argument can also be extended to the investigation of the neural consequences of multilingualism. Specifically, there is a need to characterize the key experiences that define multilingualism. In the current study, although L2AoA has been shown to be a reliable and longstanding indicator of L2 acquisition, the complexity of this variable is interwoven with other dimensions of multilingualism, such as functional usage and L2 proficiency. Indeed, in our sample, L2AoA refers to the onset age at which participants first acquired their L2 while pL2AoA accounts for proficiency and refers to the onset age at which participants first acquired their second most proficient language. In our sample, we propose that L2AoA may have a simpler meaning, referring to the starting point of which initial exposure, or the DRM's Stage 1, occurred. In studies where a positive relationship was reported, L2AoA may actually capture an experience beyond the starting point of L2 acquisition that is DRM's Stage 3, inclusive of how L2 is persistently employed across participants' lives. This also suggests that the same measure may reflect different experiences, and thus different neural consequences, in multilinguals with different immersive backgrounds.

5. Limitations and future research directions

As in any study attempting to characterize a life experience, it is challenging to balance between generalizability and specificity of the study. In the current study, we employed a convenience sample without limiting the second language spoken. Given the variety of non-English languages spoken by the sample, we were not capable of assessing participants' non-English proficiency. As a result, we were not able to examine the relationship between FA and L2 proficiency. Our intention was to examine the relationship between variation in white matter tracts and L2AoA and pL2AoA in a young adult sample. Future research could include a more homogeneous language sample in order to include language proficiency assessments. Another limitation of the current study is the small sample size. Although the sample size is comparable to other previous studies, the variation in language experience may call for a larger sample to account for variability in signal across languages. In addition, our adult sample was multilingual with the majority of the participants speaking three or more languages. Although our sample was heterogeneous, it was not sufficiently diverse enough to examine the different stages in the DRM. By including a larger sample, future research could potentially conduct subgroup analysis and examine the role of L2 immersion effects. Lastly, in terms of DTI data preprocessing, we did not acquire field maps to account for possible susceptibility to distortion on EPI sequences in 3T scanner. As a result, we were not able to perform the most current distortion correction in FSL. Without the field maps or additional acquisition with different parameters, we only performed eddy current correction.

6. Conclusion

In the present study, we have shown that two measures of age of acquisition were correlated with variation in white matter architecture in multilingual young adults. The conventional measure of L2AoA, the onset age of acquiring a second language, was shown to be negatively associated with FA in a cluster in the bilateral corpus callosum. Another measure, pL2AoA, the onset age of acquiring the *second most proficient* language, was also negatively correlated with FA in the corpus callosum, with a broader extension to bilateral corona radiata and right internal capsule. While L2AoA may be a simple and robust correlate of variation in brain structure, for adult multilinguals, this variable may encompass unmeasured complexity that calls for additional information in order to better characterize multilinguals' language use and proficiency. Finally, Pliatsikas' (2020) Dynamic Restructuring Model provides a framework for the investigation of long-term second language immersion experience. Coupled with theories on contextual language usage, such as the Adaptive Control Hypothesis (Green & Abutalebi, 2013), future research could qualify multilingual experience in terms of interaction and duration in order to provide a more comprehensive understanding of the key dimensions of multilingual experience that are relevant to brain plasticity.

Declaration of competing interest

The authors declare no conflict of interest. The authors received no funding from an external source.

Acknowledgements

This research was carried out in part at the Harvard Center for Brain Science. This work involved the use of instrumentation supported by the NIH Shared Instrumentation Grant Program; specifically, grant number S10OD020039. We acknowledge the financial support provided by the Harvard Graduate School of Education to the first author.

References

Anderson, J. A. E., Grundy, J. G., De Frutos, J., Barker, R. M., Grady, C., & Bialystok, E. (2018). Effects of bilingualism on white matter integrity in older adults. *NeuroImage*, 143–150. https://doi.org/10.1016/j.neuroimage.2017.11.038.

Anderson, J. A. E., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods, 50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9.

Andersson, J. L., Jenkinson, M., & Smith, S. (2007). *Non-linear registration aka spatial normalisation FMRIB technical report TR07JA2*. FMRIB Analysis Group of the University of Oxford.

Barratt, W. (2006). *Barratt simplified measure of social status (BSMSS)*. Indiana State University.

Boltzmann, M., Mohammadi, B., Samii, A., Münte, T. F., & Rüsseler, J. (2017). Structural changes in functionally illiterate adults after intensive training. *Neuroscience, 344*, 229–242. https://doi.org/10.1016/j.neuroscience.2016.12.049.

Cargnelutti, E., Tomasino, B., & Fabbro, F. (2019). Language brain representation in bilinguals with different age of appropriation and proficiency of the second language: A meta-analysis of functional imaging studies. *Frontiers in Human Neuroscience, 13*. https://doi.org/10.3389/fnhum.2019.00154.

Cassilhas, R. C., Tufik, S., & de Mello, M. T. (2016). Physical exercise, neuroplasticity, spatial learning and memory. *Cellular and Molecular Life Sciences: CMLS, 73*(5), 975–983. https://doi.org/10.1007/s00018-015-2102-0.

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience, 15*, 336–345. https://doi.org/10.1038/nrn3709.

Cummine, J., & Boliek, C. A. (2013). Understanding white matter integrity stability for bilinguals on language status and reading performance. *Brain Structure and Function, 218*(2), 595–601. https://doi.org/10.1007/s00429-012-0466-6.

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences of the United States of America, 116*(15), 7565–7574. https://doi.org/10.1073/pnas.1811513116.

Fricke, M., Zirnstein, M., Navarro-Torres, C., & Kroll, J. F. (2019). Bilingualism reveals fundamental variation in language processing. *Bilingualism: Language and Cognition, 22*(1), 200–207. https://doi.org/10.1017/S1366728918000482.

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. https://doi.org/10.1016/j.neuropsychologia.2013.09.037.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology (Hove, England), 25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377.

Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean, & P. Li (Eds.), *The Psycholinguistics of Bilingualism* (pp. 5–25). John Wiley & Sons.

Hämäläinen, S., Sairanen, V., Leminen, A., & Lehtonen, M. (2017). Bilingualism modulates the white matter structure of language-related pathways. *NeuroImage, 152*, 249–257. https://doi.org/10.1016/j.neuroimage.2017.02.081.

Hayakawa, S., & Marian, V. (2019). Consequences of multilingualism for neural architecture. *Behavioral and Brain Functions: BBF, 15*(1), 6. https://doi.org/10.1186/s12993-019-0157-z.

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*, 239–254. https://doi.org/10.1016/j.neuroimage.2012.06.081.

Kaufman, A. S., & Kaufman, N. L. (2004). *Manual for the kaufman brief intelligence test – second edition (KBIT-2)*. Circle Pines, MN: American Guidance Service.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences, 22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006.

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. https://doi.org/10.1016/j.bandl.2013.05.014.

Kraus, N., & White-Schwoch, T. (2017). Neurobiology of Everyday Communication: What Have We Learned From Music? *The Neuroscientist, 23*(3), 287–298. https://doi.org/10.1177/1073858416653593.

Kuhl, P. K., Stevenson, J., Corrigan, N. M., van den Bosch, J. J. F., Can, D. D., & Richards, T. (2016). Neuroimaging of the bilingual brain: Structural brain correlates of listening and speaking in a second language. *Brain and Language, 1–9*. https://doi.org/10.1016/j.bandl.2016.07.004.

Kühn, S., Gleich, T., Lorenz, R. C., Lindenberger, U., & Gallinat, J. (2014). Playing Super Mario induces structural brain plasticity: Gray matter changes resulting from training with a commercial video game. *Molecular Psychiatry, 19*(2), 265–271. https://doi.org/10.1038/mp.2013.120.

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 58*, 301–324. https://doi.org/10.1016/j.cortex.2014.05.001.

Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods, 264*, 47–56.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology (Hove, England), 25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574.

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. https://doi.org/10.1523/JNEUROSCI.4563-11.2011.

Luk, G., de Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control?*. *Bilingualism: Language and Cognition, 14*(4), 588–595. https://doi.org/10.1017/S1366728911000010.

McEwen, B. S., Nasca, C., & Gray, J. D. (2016). Stress effects on neuronal structure: Hippocampus, amygdala, and prefrontal cortex. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 41*(1), 3–23. https://doi.org/10.1038/npp.2015.171.

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*, 72–80. https://doi.org/10.1016/j.brainres.2011.12.005.

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), Article e0117968. https://doi.org/10.1371/journal.pone.0117968.

Nichols, E. S., & Joanisse, M. F. (2016). Functional activity and white matter microstructure reveal the independent effects of age of acquisition and proficiency on second-language learning. *NeuroImage*, 15–25. https://doi.org/10.1016/j.neuroimage.2016.08.053.

Pliatsikas, C. (2020). Understanding structural plasticity in the bilingual brain: The Dynamic Restructuring Model. *Bilingualism: Language and Cognition, 23*(2), 459–471. https://doi.org/10.1017/S1366728919000130.

Pliatsikas, C., & Luk, G. (2016). Executive control in bilinguals: A concise review on fMRI studies. *Bilingualism: Language and Cognition, 19*(4), 699–705. https://doi.org/10.1017/S1366728916000249.

Pliatsikas, C., Moschopoulou, E., & Saddy, J. D. (2015). The effects of bilingualism on the white matter structure of the brain. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 112*(5), 1334–1337. https://doi.org/10.1073/pnas.1414183112.

Rossi, E., Cheng, H., Kroll, J. F., Diaz, M. T., & Newman, S. D. (2017). Changes in white-matter connectivity in late second language learners: Evidence from diffusion tensor imaging. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.02040.

Schlegel, A. A., Rudelson, J. J., & Tse, P. U. (2012). White matter structure changes as adults learn a second language. *Journal of Cognitive Neuroscience, 24*(8), 1664–1670. https://doi.org/10.1162/jocn_a_00240.

Schrank, F. A., Wendling, B. J., & Alvarado, C. G. (2010). *Woodcock-Muñoz Language survey–revised normative update*. Rolling Meadows, IL: Riverside.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., … Behrens, T. E. (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., Johansen-Berg, H., … Niazy, R. K. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage, 23*, S208–S219.

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition, 22*(2), 401–415. https://doi.org/10.1017/S1366728917000682.

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*. https://doi.org/10.3389/fpsyg.2015.00164.