ELSEVIER

Cortex

CrossMark

Special section: Review

Neuroplasticity as a function of second language learning: Anatomical changes in the human brain

Ping Li[*], Jennifer Legault and Kaitlyn A. Litcofsky

Center for Brain, Behavior, and Cognition, Center for Language Science, Department of Psychology, Pennsylvania State University, United States

ARTICLE INFO

ABSTRACT

The brain has an extraordinary ability to functionally and physically change or reconfigure its structure in response to environmental stimulus, cognitive demand, or behavioral experience. This property, known as neuroplasticity, has been examined extensively in many domains. But how does neuroplasticity occur in the brain as a function of an individual's experience with a second language? It is not until recently that we have gained some understanding of this question by examining the anatomical changes as well as functional neural patterns that are induced by the learning and use of multiple languages. In this article we review emerging evidence regarding how structural neuroplasticity occurs in the brain as a result of one's bilingual experience. Our review aims at identifying the processes and mechanisms that drive experience-dependent anatomical changes, and integrating structural imaging evidence with current knowledge of functional neural plasticity of language and other cognitive skills. The evidence reviewed so far portrays a picture that is highly consistent with structural neuroplasticity observed for other domains: second language experience-induced brain changes, including increased gray matter (GM) density and white matter (WM) integrity, can be found in children, young adults, and the elderly; can occur rapidly with short-term language learning or training; and are sensitive to age, age of acquisition, proficiency or performance level, language-specific characteristics, and individual differences. We conclude with a theoretical perspective on neuroplasticity in language and bilingualism, and point to future directions for research.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

More than half of the world's population are actively learning or speaking a second language in addition to their native tongue (Grosjean & Li, 2013). What impact does experience with a second language have on the human brain? Traditionally, both folk wisdom and scientific evidence point to the decreasing plasticity of the adult brain in acquiring a new language, especially given the arguments of the so-called "critical period hypothesis" (Kennedy & Norman, 2005;

Lenneberg, 1967). Recent scientific evidence, however, has challenged this view. In particular, cognitive and brain studies of bilingual language acquisition, along with studies of memory, attention, and perception, have demonstrated continued neuroplasticity for language learning in the adult brain that has never been previously imagined (see Abutalebi & Green, 2007; Hernandez, 2013; Li, 2014, for reviews). The study of neuroplasticity of language learning in adulthood, along with the understanding of neural correlates of language processing and representation, has made significant progress in the last decade thanks to rapid advances in neuroimaging technologies (see reviews in Hickok, 2009; Poeppel, Emmorey, Hickok, & Pylkkänen, 2012; Price, 2000, 2010; Richardson & Price, 2009; Rodriguez-Fornells Cunillera, Mestres-Misse, & de Diego-Balaguer, 2009).

Experience-dependent neural changes can result from many aspects of environmental input, cognitive demand, or behavioral experience, but the intensity and frequency of language use may be particularly powerful in bringing about such changes in the brain (see Bates, 1999 for an earlier synthesis; see Bialystok & Barac, 2013 for a recent discussion). Many people are born bilingual in our increasingly more connected and multilingual world, while many others are learning a new language later in life due to travel, business, or immigration. Globalization, widespread use of digital technology, and increased cross-cultural communication provide further impetus to the rapid rise of bilingualism and multilingual societies. The study of the bilingual brain originally arose from neuroscientists' interest in understanding how the same brain supports and represents two or more languages. Since the mid-to-late 1990s, a large number of neurocognitive studies, using neuroimaging methods such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and electroencephalography/event-related potential (EEG/ERP), have revealed specific functional brain patterns in the learning of a second language (L2) (see reviews in Abutalebi, Cappa, & Perani, 2005; Costa & Sebastián-Gallés, 2014; Hernandez, 2013; Indefrey, 2006; Li & Tokowicz, 2012; van Hell & Tokowicz, 2010). These studies indicate that in contrast to predictions of the critical period hypothesis, L2 learning, even if it occurs late in adulthood, lead to both behavioral and neural changes that may approximate the patterns of native or first language (L1).

Even more surprising is that the neural patterns of L2 experience are often, if not always, accompanied by anatomical changes in brain structure. Such anatomical changes can occur in the form of, for example, increased gray matter (GM) density, increased cortical thickness (CT), or enhanced white matter (WM) integrity. A number of recent studies have thus begun to examine the structural or anatomical changes induced by L2 experience on the brain. Given the significant anatomical changes that have been reported for memory, attention, and other cognitive domains (see Section 3.2), it is important that we consider the anatomical substrates of second language learning. In this article, we provide an overview and synthesis of the relevant studies, and identify key variables and mechanisms underlying language experience related structural neuroplasticity. We aim at not only reviewing the emerging literature, but also identifying the common principles that drive brain changes in order to integrate our knowledge of the structure-function-behavior relationships.

2. Anatomical correlates of second language learning

Research in bilingualism and second language[1] has generated much enthusiasm lately in the study of the mind and the brain (see Diamond, 2010). What has brought bilingualism to the spotlight? There may be several reasons but one key line of research behind the current enthusiasm is the neurocognitive impact that the learning and use of multiple languages may have on the brain (see reviews in Bialystok, 2009; Costa & Sebastián-Gallés, 2014; Hernandez, 2013; Li, 2014). The bilingual brain is a highly adaptive system, and it responds to multiple language experiences flexibly and reflects the adaptive dynamics as both functional and anatomical brain changes. In this section we review the major evidence that has accumulated in the last decade on how the learning of L2, or bilingual experience more generally, may bring about anatomical changes in the brain.

Functional neuroimaging methods, especially fMRI, have played a key role in the study of bilingualism and second language acquisition (see Abutalebi, Cappa, & Perani, 2005; Grosjean & Li, 2013, Chapter 10; Hernandez, 2013; Indefrey, 2006; Li & Tokowicz, 2012; for reviews). While functional neuroimaging has led to a significant understanding of the bilingual brain, the use of structural imaging techniques has only begun recently in the study of bilingualism and second language. As we will discuss below, structural imaging methods allow us to measure brain changes in anatomical structure and may offer broader implications for understanding the bilingual brain, particularly with regard to their ability to identify causal links between experience and neuroplasticity through training. Let us first briefly review the three major measures and the methodologies with which we can identify learning-induced or experience-dependent changes in the brain's anatomical structures.

2.1. Measures of anatomical changes

Neurons are organized within the brain to form GM and WM. GM consists primarily of neuronal cell bodies, whereas WM consists of axons and support cells (e.g., glia cells). Bundles of axons form the so-called fiber tracts that connect different cortical regions within the same hemisphere (through association tracts), between hemispheres (through commissures, e.g., the corpus callosum (CC)), or between cortical and subcortical structures (projection tracts). The brain is filled with cerebrospinal fluid (CSF), which also runs through the ventricles of the brain. Measures of anatomical changes focus mainly on changes in GM and WM.

---

[1] Many people learn or speak a third or fourth language. Here we use "bilingualism" or "second language" as a generic and inclusive term to cover situations of two or *more than* two languages.

2.1.1. GM density

GM density or volume has been one of the most common measures of anatomical brain changes. Although it is not entirely clear what exactly an increase in GM volume entails at a microstructural level, it is generally believed that it reflects an aggregate measure of the changes in cell size of both neurons and glial cells, neurogenesis associated with both neurons and glial cells, and possible changes in the intra-cortical axonal architecture including synaptogenesis (May & Gaser, 2006; Zatorre, Fields, & Johansen-Berg, 2012). Thus, GM density does not directly translate to the density of neurons, or other simple measure of the brain morphology. To identify GM density, researchers rely on voxel-based morphometry (VBM), an analytic method that extracts GM information from structural MRI scans (see Ashburner & Friston, 2000; Mechelli, Price, Friston, & Ashburner, 2005). VBM typically involves the normalization of each brain scan to a standard stereotactic space (e.g., MNI space), delineation of gray versus WM versus CSF, and a voxel-by-voxel analysis of the tissue concentration. VBM identifies the local tissue environment after correction for macroscopic anatomical differences across participants.

2.1.2. Cortical thickness

CT, also based on structural MRI scans, measures the thickness of GM (Fischl & Dale, 2000; Kim et al., 2005; Lerch & Evans, 2005). Unlike GM density or volume, CT is a direct measure of cortical morphology. In this technique, voxels are first segmented into GM, WM, or CSF. The boundaries between GM and WM, and between GM the pia mater are then delineated either manually or through automated procedures. Finally, the thickness between these surfaces is measured using a variety of methods, each determining the distance between nodes on each surface for the entirety of the cortex examined. CT provides sub-millimeter accuracy and takes into account the folding of the cortical surface. Structurally there may be an inverse relationship between CT and GM due to the cortical folding patterns: thicker cortical regions are less convoluted and therefore have less GM density (see Chung, Dalton, Shen, Evans, & Davidson, 2006). It is relatively insensitive to differences in MRI scanners and parameters, but is less accurate for areas where the GM/WM boundary is less clear, such as in primary sensory areas that contain more myelination.

2.1.3. WM integrity

WM integrity refers to a measure based on data from diffusion tensor imaging (DTI), a technique that examines the diffusion of water molecules in the brain. DTI compares the degree of diffusivity of neurons along the axon, referred to as axial diffusivity (AD) along with the radial diffusivity (RD) that is perpendicular to the axon diameter (Filler, 2009). Another measure, the mean diffusivity (MD), is used to measure diffusion within a voxel, regardless of orientation, and is calculated by averaging the eigenvalues (Alexander, Lee, Lazar, & Field, 2007). Lower MD values often correspond to greater WM integrity. By far the most commonly used value to calculate the magnitude of diffusion is the fractional anisotropy (FA), a normalized standard diffusivity value between 0 and 1 calculated from the eigenvectors of the diffusion tensor (Assaf & Pasternak, 2008). FA has been a yardstick of

WM integrity in the literature, where a value of 0 indicates an isotropic environment as is seen in the ventricles of the brain, and a value of .2–1 an anisotropic environment as is seen in WM tracts (Kunimatsu et al., 2004). The higher the FA value, the more integrity the WM has (contrasting the interpretation of the MD value). In addition, a high FA value, when coinciding with a low RD value, could suggest increased myelination.

A sizable number of studies in the last decade have used the above three methods to examine anatomical changes in the brain as a function of bilingual or L2 experience (see Richardson & Price, 2009 for a review of monolingual studies using these methods). Table 1 presents a summary of these studies and Table 2 an overview of the different regions and tracts of interest modulated by bilingual experience. Fig. 1 presents a direct comparison between bilinguals and monolinguals in the areas where anatomical changes have been observed, and how these changes correlate with behavioral tasks or variables. We discuss the details of these studies in the sections that follow.

2.2. Structural brain changes induced by bilingual experience in children and adults

One of the pioneering studies using VBM to examine GM density in bilingual learners was Mechelli et al. (2004). In this study, bilinguals were participants who had learned a European language before the age of 5 (early bilinguals) or between the ages of 10 and 15 (late bilinguals). In general, the bilinguals showed greater GM density in the left inferior parietal lobule (IPL)[2] than did monolinguals, but the effect was greater in the early bilinguals than in the late bilinguals. The IPL has been previously implicated in functional imaging studies as an important area for phonological working memory, lexical learning, and semantic integration (Baddeley, 2003; Della Rosa et al., 2013; Mechelli et al., 2004;). The expansion of this area might be particularly related to the bilingual's acquisition and processing of a larger vocabulary due to the L2 (see Richardson & Price, 2009). More important, Mechelli et al. also showed that the extent of GM density increases was positively associated with the proficiency of the learner in the L2 (more proficient, more GM), and negatively correlated with the learner's age of L2 acquisition (the earlier the learning, the more the GM).

Other studies have since replicated Mechelli et al.'s finding and confirmed the role of the IPL and adjacent regions in the temporo-parietal cortex for bilingualism, showing that bilinguals, in general, have greater GM density than monolinguals in this brain area (e.g., Della Rosa et al., 2013; Grogan et al., 2012; see discussion in Section 2.3). More specifically, IPL, including the posterior supramarginal gyrus (SMG), has been implicated to play an important role for vocabulary knowledge in general, for both L1 and L2: Lee et al. (2007) found a significant positive correlation between monolingual vocabulary size and GM volume in the bilateral IPL and posterior SMG. Additionally, Xiang et al. (2012) showed how variability in structural pathways was related to language abilities as

---

[2] See Appendix for all abbreviations of the brain regions used in this paper. We follow accepted conventions in the literature for most if not all acronyms.

Table 1 – Studies of structural changes associated with language experience and short-term training.

Study	Groups	Age	Mean AoA	MRI methods	GM/WM differences	Brain-behavior correlations[a]
Abutalebi et al., 2012	17 German-Italian BI 14 Italian MO	23.4 26.6	~5 N/A	er-fMRI VBM	−	*Flanker; Language Switching* BI:(+) GM density ACC & functional conflict effect (−) GM density ACC & behavioral conflict effect
Abutalebi et al., 2014	23 Chinese BI (12 L1 Cantonese, L2 English, 11 two Chinese dialects) 22 Italian MO	62.2 61.9	18.9 N/A	VBM	BI > MO: L ATP	*L1 & L2 picture naming* (+) GM volume in L ATP & L2 picture naming
Cummine & Boliek, 2013	12 Chinese-English BI 11 English MO	24.2 28.5	>5 N/A	DTI	MO > BI: R IFOF, R superior ATR, bilateral inferior ATR	*Word reading:* BI: (−) R ITG, L EC, L CN, etc. & RT to *inconsistent* words (−) L SFG, R CB, R STG, L lateral sulcus, etc. & RT to *consistent* words MO:(−) L SFG & R IFG, R SN, L IPL, L POS & RT to *inconsistent* words (−) cingulate sulcus, etc. & RT to *consistent* words
Della Rosa et al., 2013	15 ML Longitudinal design, scanned twice, separated by 1 year	9.9	Early	VBM	−	*Attentional Network Task (ANT)* Change over time: (+) IPL & multilingual talent interaction
Elmer et al., 2011	12 SI 12 Controls Language not reported	37.9 28.4	Not given	DTI	Controls > SI: R IPL, dorsal R CN, among others	−
Elmer et al., 2014	12 SI 12 ML Varied languages	37.0 27.4	Varied, Late	VBM	ML > SI: L SMG, bilateral IFGpt, L IFGop, among others	SI: (−) R IFGop, L IFGpt, bilateral CN, middle-anterior cingulate gyrus & cumulative number of interpretation practice hours
García-Pentón et al., 2014	13 Spanish-Basque BI 13 Spanish MO	24.1 29.1	.5 N/A	DTI Network analysis	BI > MO: More connected sub-network: L insula, STG, IFGop, IFGpt, SMG, medial SFG More connected sub-network: L SOG, L STP, L AG, L SPG, R SFG Higher global efficiency in both sub-networks MO > BI: Higher global efficiency in whole network	−
Gold et al., 2013	20 BI 63 MO English one language for all	63.9 64.4	<10 N/A	VBM DTI	GM: WM: FA: MO > BI: ILF/IFOF, fornix, portions of CC RD: MO < BI: ILF/IFOF, fornix, portions of CC	−
Grogan et al., 2012	30 BI 31 ML English is non-native language for all	26.7 26.9	8.2 English = 6.4	VBM	ML > BI: R posterior SMG	*Phonemic fluency; Lexical decision task (LDT)* BI: (+) L IFGop & phonemic fluency, LDT (−) L IFGop & AoA
Hosoda et al., 2013	137 Japanese-English learners	24.0	11	VBM DTI	−	*English Vocabulary Test (EVT); National Adult Reading Test (NART)* GM: (+) IFGop, CN, STG/SMG, ACC (all bilateral) & EVT WM: (+) Connectivity of R IFGop-CN, etc. & EVT (+) FA in R IFGop, R ILF, R AF & EVT
	24 Japanese-English learners, (TG) 20 Japanese-English learners, (CG) *Training*[b]: 4-month laboratory training on L2 vocabulary	20.1 20.1	11 11	VBM DTI	GM: TG > CG: R IFGop WM: TG > CG: R IFGop IFGop-caudate & R dorsal pathway connectivity	*Test of English for International Communication (TOEIC); EVT; NART* GM & WM: (+) R IFGop & TOEIC (+) R IFGop-caudate & TOEIC

Table 1 – (continued)

Study	Groups	Age	Mean AoA	MRI methods	GM/WM differences	Brain-behavior correlations[a]
Klein et al., 2013	12 BI_{SIM} 25 BI_E 29 BI_L 22 MO	23 26 28 25	1 5 10 N/A	CT	BI_E, BI_L > MO: L IFG BI_L < BI_E, BI_{SIM}, MO: R IFG BI_E < MO: R IFG	*L2 proficiency: subjective ratings & objective screening* BI_E < MO: (+) L IFG, Left SPL & AoA (−) R IFG & AoA
Kwok et al., 2011	19 MO *Training*: 2-hour laboratory training on new color words	Adult	N/A	VBM	GM: V2/3	
Luk et al., 2011	14 BI English L1 or L2 14 English MO	70.5 across groups	<11 N/A	DTI rs-FC	BI > MO: CC stretching to bilateral SLF, R IFOF, & uncinate fasciculus	(+) WM connectivity rs-FC, especially in frontal regions
Mårtensson et al., 2012	14 SI TG 17 CG All native Swedish *Training*: 3-month intensive language course focused on vocabulary	19.9 20.6	N/A N/A	CT HP volume	*SI TG* > *CG*: L MFG, L IFG, L STG, R HP	*Grades on translation & interpretation exam 2 months after post-scan* SI TG > CG: (+) L STG, R HP & L2 proficiency (+) L MFG & struggle
Mechelli et al., 2004	25 BI_E 25 BI_L 25 MO All native English	Not given	<5 10-15 N/A	VBM	BI > MO: L IPL BI_E > BI_L: L IPL, R IPL	−
	22 Italian-English BI	Not given	2–34	VBM	−	*L2 proficiency: reading, writing, speech production & comprehension* (+) L IPL & L2 proficiency (−) L IPL & AoA
Mohades et al., 2012	15 BI_{SIM} 15 BI_{SE} 10 MO L1: French or Dutch	9.3 9.7 9.6	<3 >3 N/A	DTI	BI_{SIM} > BI_{SE} > MO: IFOF MO > BI_{SE} > BI_{SIM}: AC-OL	−
Pliatsikas et al., 2013	17 Greek learners of English 22 English MO	27.5 24.5	7.7 N/A	VBM	BI > MO: bilateral CB	*English regular/irregular past tense; lexical decision task* BI: (−) Bilateral CB & RT to regular condition (+) L & R posterior putamen & years of exposure
Ressel et al., 2012	22 Spanish-Catalan BI 22 Spanish MO	21.5 23.1	<7 N/A	VBM HG volume	BI > MO: HG volume; GM volume L HG (not WM)	−
Schlegel et al., 2012	11 English MO TG 16 English MO CG *Training*: 3-semester intensive university Chinese course	20.05 across groups	N/A N/A	DTI	*TG* > *CG*: FA: Frontal tracts that cross the genu of the CC *CG* > *TG*: RD: Frontal tracts that cross the genu of the CC	*L2 proficiency: instructor evaluation* TG > CG: (+) FA & amount of Chinese learned
Stein et al., 2012	10 English MO *Training*: 5-month study abroad in Switzerland learning German	17.5	N/A	VBM	L IFG, L ATL	(+) L IFG, L ATL & L2 proficiency
Wong et al., 2008	17 English MO *Training*: Weeks-long laboratory training on using lexical tone for word identification	21	N/A	HG volume	*More* > *less successful learners*: GM L HG (trend WM)	*Task of mapping words to linguistic pitches* (+) GM, WM HG & task performance
Zou et al., 2012	14 Chinese-sign language BI 13 Chinese MO	49 48	29 N/A	er-fMRI VBM	BI > MO: L CN	*Language switching task* BI > MO: (+) L CN GM & L CN activation during language switching task

Abbreviations: AoA: age of acquisition; BI: bilinguals; BI_{SIM}: simultaneous bilinguals; BI_{SE}: sequential bilinguals; BI_E: early bilinguals; BI_L: late bilinguals; CG: control group; er-fMRI: event-related functional MRI; ML: multilinguals; MO: monolinguals; N/A: not applicable; rs-FC: resting-state functional connectivity; RT: reaction time; SI: simultaneous interpreters; TG: training group. 'Not given' indicates that the information was not provided in the original paper. See Appendix for abbreviation keys to brain regions.

[a] Behavioral tasks are shown in italics and correlations are denoted with '(+)' for positive relationships and '(−)' for negative relationships.

[b] For training studies, main findings indicate changes from pre- to post-training.

Table 2 – Regions and tracts of interest associated with structural and behavioral changes due to language experience.

ROIs and TOIs[a]	Structural changes in response to language experience	Behavioral correlations[b]
Training Studies		
L ATL	GM: TG (Stein et al., 2012)	GM: (+) Language Proficiency (Stein et al., 2012)
	TG > CG (Hosoda et al., 2013)	(+) L2 proficiency (Hosoda et al., 2013)
CN		GM: (+) Phonemic Fluency (Grogan et al., 2009)
L IFG	GM: TG (Stein et al., 2012)	GM: (+) Language Proficiency (Stein et al., 2012)
	CT: TG > CG (Mårtensson et al., 2012)	
HP	GM: TG > CG (Mårtensson et al., 2012)	GM: (+) L2 proficiency (Mårtensson et al., 2012)
L HG	GM: TG (Wong et al., 2008)	GM: (+) Linguistic pitch-word identification, (−) RT (Wong et al., 2008)
L ITC		GM: (+) Semantic Fluency (Grogan et al., 2009)
L MFG	CT: TG > CG (Mårtensson et al., 2012)	CT: (+) Struggle (Mårtensson et al., 2012)
preSMA		GM: (+) Phonemic Fluency (Grogan et al., 2009)
L STG	CT: TG > CG (Mårtensson et al., 2012)	CT: (+) L2 proficiency (Mårtensson et al., 2012)
V2/3	GM: TG (Kwok et al., 2011)	–
CC	FA: TG > CG (Schlegel et al., 2012)	FA: (+) L2 proficiency (Schlegel et al., 2012)
Non-Training Studies		
ACC		GM: (+) functional conflict effect, (−) behavioral conflict effect (Abutalebi et al., 2012)
CB	GM: TG > CG (Pliatsikas et al., 2013)	GM: (+) Regular Verb Morphology Processing (Pliatsikas et al., 2013)
L HG	GM: BI > MO (Ressel et al., 2012)	–
L IFG	CT: BI > MO (Klein et al., 2013)	CT: (+) AoA (Klein et al., 2013)
	GM: BI > ML (Grogan et al., 2012)	GM: (+) Lexical decision, (+) Verbal fluency, (−) AoA (Grogan et al., 2012).
R IFG	CT: BI_E and BI_{SIM} > BI_L	CT: (+) AoA (Klein et al., 2013)
	MO > BI (Klein et al., 2013)	
L IPL	GM: BI > MO,	GM: (+) L2 Proficiency, (−) AoA (Mechelli et al., 2004)
	BI_E > BI_L (Mechelli et al., 2004)	(+) Multilingual talent interaction (Della Rosa et al., 2013)
SMG	GM: ML > BI (Grogan et al., 2012)	
L ATL	GM: BI > MO (Abutalebi et al., 2014)	GM: (+) L2 Picture Naming Task (Abutalebi et al., 2014)
AC-OL	FA: MO > BI_{SE} > BI_{SIM} (Mohades et al., 2012)	–
CC	FA: BI > MO (Luk et al., 2011)	–
	MO > BI (Gold et al., 2013)	
FX	FA: MO > BI (Gold et al., 2013)	–
L IFOF	FA: MO > BI (Gold et al., 2013)	–
	BI_{SIM} > BI_{SE} > MO (Mohades et al., 2012)	
R IFOF	FA: MO > BI (Cummine & Boliek, 2013)	–
	MO > BI (Gold et al., 2013)	
	BI > MO (Luk et al., 2011)	
ILF	FA: MO > BI (Gold et al., 2013)	–
SLF	FA: BI > MO (Luk et al., 2011)	–
UF	FA: BI > MO (Luk et al., 2011)	–

Abbreviations: AoA: age of acquisition; BI: bilinguals; BI_E: early bilinguals; BI_L: late bilinguals; BI_{SE}: sequential bilinguals; BI_{SIM}: simultaneous bilinguals; CG: control group; CT: cortical thickness; FA: fractional anisotropy; GM: gray matter volume; MO: monolinguals; ROIs: regions of interest; TG: training group; TOIs: tracts of interest. See Appendix for abbreviation keys to brain regions.
[a] ROIs are listed in alphabetical order followed by TOIs in alphabetical order.
[b] Correlations are denoted with '(+)' for positive relationships and '(−)' for negative relationships. For training studies, main findings indicate changes from pre- to post-training.

measured by a standardized language aptitude test. They found a significant relationship between performance on rapid vocabulary learning and the pathway connecting inferior frontal gyrus (IFG, Brodmann's Area or BA 47) and the parietal lobe (see also discussion in Section 3.5). In another study, Klein, Mok, Chen, and Watkins (2013) measured GM thickness as CT, and found that in the left IFG as well as superior parietal lobule (SPL), CT correlated with the learner's age of onset for L2 acquisition (L2 AoA) in the opposite direction: the later the learning, the greater the CT. Interestingly, although this appears to contrast with patterns found in most other studies (where a negative correlation was found between AoA and GM volume/density), these CT findings actually are consistent with the literature of GM studies, given the expected inverse

relationship between CT and GM due to cortical folding patterns (see discussion in Section 2.1.2).

In addition to GM density changes as a function of bilingual experience, WM integrity has also been shown to differ between bilinguals and monolinguals. Cummine and Boliek (2013) found higher FA values for adult monolinguals over bilinguals in the right inferior fronto-occipital fasciculus (IFOF) and the anterior thalamic radiation (ATR), which the authors interpreted as to reflect immature WM for the bilinguals (young adults in early 20s) versus the monolinguals (older adults in late 20s). Interestingly, Mohades et al. (2012) demonstrated that these differences between monolinguals and bilinguals may be tract-dependent. Their study examined WM integrity in the left hemisphere (LH) and CC in
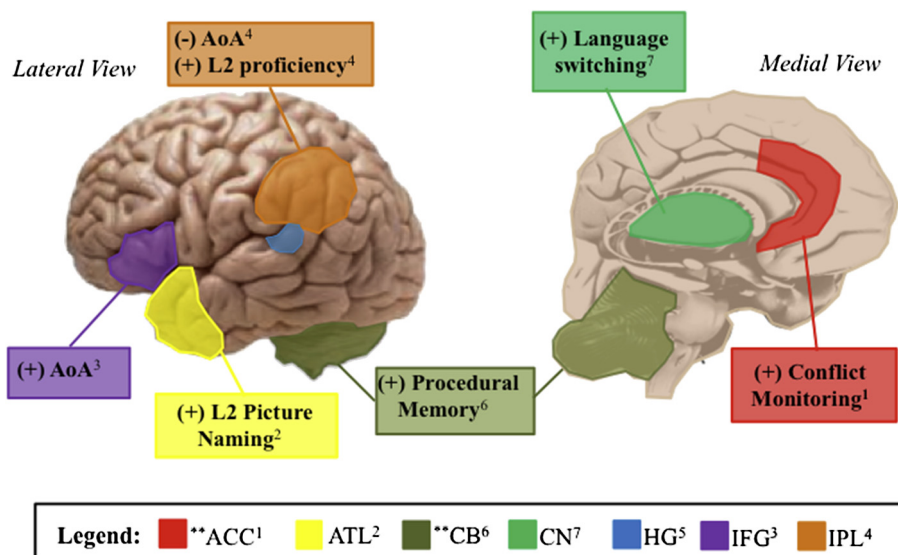
Fig. 1 – Regions that show increased GM density/volume or CT according to a number of studies with group comparisons of bilinguals versus monolinguals. Regions labeled with ** in the legend indicate bilateral GM; otherwise localized in the left hemisphere or medial section. Further, structural increases in the specific regions are shown to be correlated with behavioral tasks or variables: (+): positive correlation with a task or variable; (−): negative correlation with a task or variable. These brain-behavior correlations are based on the following studies: [1]Abutalebi et al., (2012); [2]Abutalebi et al., (2014); [3]Klein et al., (2013); [4]Mechelli et al., (2004); [5]Ressel et al., (2012); [6]Pliatsikas et al., (2013); [7] Zou et al., (2012). See Appendix for abbreviation keys.

bilingual and monolingual children who had learned their L2 either simultaneously or sequentially. Their results indicated that bilingual children have greater FA values in the left IFOF than their age-matched monolinguals, which, when compared with the Cummine and Boliek study, highlights the differences in WM integrity between the right and left hemispheres. These data show that structural properties of the brain such as WM can be modified from early on with L2 experience. The WM changes also reflected an incremental effect: increases in IFOF depended on whether L2 occurred early or late in life. On first appearance, the patterns from Cummine and Boliek's and Mohades et al.'s studies seem to be contradictory, but it is important to note that (a) developmental trajectories in WM integrity might differ between tracts, as mentioned; for example, Mohades et al. also found lower FA values in bilinguals as compared to monolinguals in the fiber bundles arising from the anterior CC to the orbital lobe (AC-OL), and (b) there are age effects: the Cummine and Boliek's study involved older participants in their study: age 24 (bilinguals) and age 29 (monolinguals), while the Mohades et al. study involved 9 year olds for both the bilingual and monolingual groups. Cummine and Boliek suggested that it may be difficult to see structural brain differences in mid-adulthood as a function of status of bilingualism.

Given the different patterns reported for children versus adults, what happens to the elderly? In Section 3.4, we will further discuss the age effect and its implications (see also a recent review by Antoniou, Gunasekera, & Wong, 2013). Here we focus on a few recent studies that have examined aging adults. Luk, Bialystok, Craik, and Grady (2011) assessed older bilinguals and monolinguals at a mean age of 70 who spoke

English as their first language. All bilinguals had begun to learn their second language before age 11 and used both languages regularly throughout their lives. The authors examined both WM integrity and resting-state functional connectivity. Bilinguals, as compared to monolinguals, showed higher FA values in the CC projecting to the bilateral superior longitudinal fasciculi (SLF), the right IFOF, and the uncinate fasciculus (UF), consistent with the findings from bilingual children in Mohades et al. (2012). In addition, WM structural connectivity correlated with resting-state functional connectivity, particularly for the frontal regions, indicating parallel structure-function changes as a result of lifelong bilingualism (see further discussion in Section 3.1 regarding structure-function correspondences). Highlighting these patterns further, the authors suggested that the GM atrophy seen in patients with Alzheimer's Disease may be compensated for by increased WM integrity in bilinguals, which provides the neural basis for the idea of "cognitive reserve": lifelong bilingual experience may serve as a major deterrent to the onset of age-related cognitive decline.

The proposal that bilingualism provides a protective mechanism against age-related cognitive decline was further examined in a recent study by Abutalebi et al. (2014). These authors examined one group of older participants from Hong Kong who spoke either Chinese and English or Cantonese and Mandarin as bilinguals, and compared their brain structures and language performance with a control group of monolingual Italian speakers in a picture naming task in L1 (monolinguals and bilinguals) and L2 (bilinguals only). The two groups were matched on age (mean age of 62), education, and cognitive abilities. The structural imaging data indicated that bilinguals had greater GM volume than monolinguals,

especially in the left anterior portion of the temporal pole (TP). Further ROI-based analyses also showed that L2 naming but not L1 naming performance was positively correlated with GM volume in the left TP. The authors suggested that the TP might play an important role in bilingual lexical conceptual processing, and that bilingual experience serves as a protective factor to the rapid decrease of GM volume in this age range for older adults.

A third study of the aging population showed somewhat different patterns from the above two: Gold, Johnson, and Powell (2013) found greater FA values for aging monolinguals than for bilinguals in the bilateral inferior longitudinal fasciculus (ILF), IFOF, fornix (FX), and many portions of the CC. In fact, their WM measures for the bilinguals — reduced FA and increased RD — are indicative of neurodegenerative patterns linked to mild cognitive impairment. These authors argued differently from Luk et al. (2011), suggesting that bilinguals do not necessarily have increased WM, but may actually compensate for their WM degeneration through the use of executive function networks that are shown to be more efficient in bilinguals. The discrepancy between the studies demands more structural as well as functional imaging data for a better understanding of the nature of cognitive reserve. Gold et al. also conducted VBM analyses of their data but found no significant differences in GM volume between bilinguals and monolinguals, which seemed to suggest that GM and WM changes may occur independently (no GM but WM changes, or vice versa; see further discussion in Section 3.1 on this point).

2.3. Structural brain changes as a result of cognitive control experience

Second language experience-induced neural changes have been previously studied in connection with the bilingual cognitive advantage hypothesis (see Bialystok, 2009; Bialystok & Barac, 2013; Dong & Li, 2014; for reviews). Briefly, Bialystok and colleagues have argued that bilingualism confers distinct advantages in cognitive control, according to which bilinguals develop better executive functions (e.g., inhibiting, updating, switching) and better conflict monitoring abilities than monolinguals, in tasks that involve selective attention, inhibition of irrelevant information, and task switching. The source of bilingual cognitive advantages has been discussed in terms of the bilingual's lifelong experience in monitoring, switching between, and selecting among competing languages. The neural basis of this advantage has also been examined, including the idea of "cognitive reserve" discussed above according to Abutalebi et al. (2014), Gold et al. (2013), and Luk et al. (2011). In this context, previous studies have shown functional neural activities (increases or decreases) in several cognitive control areas, including the left IFG, the anterior cingulate cortex (ACC), the IPL, and subcortical regions including the basal ganglia (BG), particularly the left caudate and the putamen (see Abutalebi, 2008; Abutalebi & Green, 2007 for reviews). These regions form an integrated network for bilingual control, with the ACC playing a distinct role in conflict monitoring and attention (Abutalebi et al., 2012; Li, Yang, Scherf, & Li, 2013). How do structural brain changes reflect the effects of enhanced cognitive control abilities due to lifelong bilingual experience, if such experience does confer distinct

advantages in executive functions? A number of recent studies have revealed anatomical correlates of cognitive control in the bilingual brain as compared with the monolingual brain. These studies show anatomical changes as increased GM density in the brain's critical control regions. For example, Abutalebi et al. (2012) assessed whether elements of bilingual advantage such as conflict monitoring are specific to language learning or if these skills are generalizable to other domains. The authors combined an event-related (er-fMRI) design with VBM, to examined correlations between GM density and functional brain activation and also with behavioral performance. Their results, as shown in Fig. 1, indicated that in the ACC the GM volume was positively correlated with functional activity (reflecting conflict monitoring), and this correlation was stronger for bilinguals than for monolinguals (not statistically significant for the latter). However, the ACC activation was lower in bilinguals than in monolinguals when the two groups were directly compared, indicating that bilinguals need less activity in the ACC for effective conflict monitoring. In addition, the authors observed a negative correlation between GM and behavioral conflict effect, suggesting that the greater the GM volume, the smaller the conflict that bilinguals experience. Overall, these data suggest that bilinguals are more efficient at conflict monitoring and in cognitive control in general.

Della Rosa et al. (2013) further tracked fifteen bilingual children in a longitudinal design, and compared the participants' GM volume at two time points (T1 and T2) spanning a year. They also measured the bilinguals' overall language competence and their cognitive control abilities in order to relate brain changes to behavioral performance. Specifically, they calculated an interaction effect between language competence and cognitive control to predict whether changes in GM density varied as a function of development due to bilingual language experience. Their results indicated a strong positive correlation between this interaction effect and the GM density change (T2–T1) clearly in the IPL, while an inverse relationship existed between language competence and cognitive control (better language ability, less demand on control). Thus, they proposed that the cognitive demands instigated by growing up with additional languages can lead to greater GM density in the IPL, a region that "houses" both the multilingual talent and the cognitive control ability (see also Mechelli et al., 2004; Richardson & Price, 2009).

Other corroborating studies have shown that bilinguals have greater GM volume in several other areas, including the caudate nucleus (CN, Grogan, Green, Ali, Crinion, & Price, 2009; Zou, Ding, Abutalebi, Shu, & Peng, 2012) and the bilateral cerebellum (CB; Pliatsikas, Johnstone, & Marinis, 2013; see discussion below), as also shown in Fig. 1. For example, Grogan et al. (2009) showed that increased GM density in the CN, particularly bilateral head of CN, is associated with better performance in a phonemic fluency task in the L2 for high-proficiency bilinguals. GM density in the left inferior temporal cortex, however, correlated only with the semantic fluency task, similarly for both L1 and L2. These data are consistent with the idea that CN plays an important role in detecting phonological anomalies, in the selection of competing verbal responses and in language switching based on previous neuroimaging findings and patient studies (e.g., Abutalebi, 2008; Crinion et al., 2006; Zou et al., 2012).

Activation of the CN may also be reflected in the need to inhibit the L1 while performing L2 production, especially when phonological detection is crucial to the task as in phonemic fluency.

Likewise, WM studies have found greater involvement of the control network (and sub-networks) for bilinguals as compared to monolinguals. Garcia-Penton, Perez, Iturria-Medina, Gillon-Dowens, and Carreiras (2014) employed DTI analyses combined with approaches of network-based statistics (NBS) and graph theory to assess multiple connectivity differences between monolingual and bilingual adults. The advantages of using NBS approaches include the ability to examine functional and structural connections (edges) between key regions (nodes) as well as their interaction within and between sub-networks (modules), providing a more informative method of discerning mechanisms underlying cognition (Bressler & Menon, 2010). Results from this NBS approach identified two sub-networks with higher interconnectedness specific to bilinguals. The first sub-network consisted of areas in the LH, including the insula, superior temporal gyrus (STG), SMG, IFG, and medial superior frontal gyrus (mSFG), areas which have been implicated in various aspects of language processing, attention, and control. Many of these areas also show GM density increases for bilingual learners (see also Grogan et al., 2012; Klein et al., 2013; Mårtensson et al., 2012). The second sub-network consisted of regions including the left superior occipital gyrus (SOG), right superior frontal gyrus (SFG), left superior parietal gyrus (SPG), left superior temporal pole (STP) and the left angular gyrus (AG), areas that the authors posit as possible contributors to word recognition, reading, and lexical semantic processing. Garcia-Penton et al. suggested that the increased degree of interconnectivity is specific to bilinguals due to additional language experience. The fact that both of these sub-networks included a frontal region also confirms the important role that the frontal cortex plays in both language processing and cognitive control (Abutalebi & Green, 2007; Gabrieli, Poldrack, & Desmond, 1998).

2.4. Structural brain changes with regard to age of learning and proficiency in the L2

In addition to group differences between bilinguals and monolinguals, researchers have also investigated whether different L2 AoA would differentially affect neuroanatomical substrates observed. As reviewed in Section 2.2, the Mechelli et al. (2004) study revealed a strong relationship between L2 AoA and GM density in the left IPL, suggesting that the earlier the bilingual begins L2 learning, the more GM density the bilingual has. Klein et al. (2013) also showed the effect of AoA in the CT of the IFG and SPL, and Mohades et al. (2012) provided evidence of WM correlates of AoA (see Section 2.2). Overall, it appears that learning a second language early in life may result in increased anatomical changes for language related areas, in both GM density and WM integrity. It should be noted, however, that most of the studies examining the AoA effect were based on cross-sectional data from bilinguals with mixed language backgrounds, and in future studies it would be important to conduct longitudinal studies to track anatomical changes in the same individuals (e.g., see Schlegel, Rudelson, & Tse, 2012; discussed in Section 2.5).

Previous functional imaging studies may have confounded AoA with the bilingual's level of proficiency in the L2 (e.g., Kim, Relkin, Lee, & Hirsch, 1997; see further discussion later on this point), and researchers have since attempted to separate the role of AoA from proficiency and identify their independent contributions to the functional neural patterns observed (see Abutalebi & Green, 2007; Wartenburger et al., 2003). Similarly, a number of structural imaging studies have revealed anatomical changes associated with proficiency in addition to AoA. For example, Mechelli et al. (2004) showed that while AoA negatively correlated with GM density in both the left and right IPL, L2 proficiency positively correlated with GM density in the left IPL. Grogan et al. (2012) showed that performance levels in L2 lexical decision and verbal fluency tasks positively correlated with GM density in the left IFG *pars opercularis* (IFGop) for young adult bilinguals, although surprisingly not for multilinguals (those who spoke at least 3 languages). The authors suggested that the lack of this correlation for multilinguals may be because only one language (English) was tested, which may not capture the full variability in the lexical performance across all the languages of the multilingual participants'.

In another study, Zou et al. (2012) showed that bimodal bilinguals had greater GM volume in the left CN, as compared with monolinguals. The authors posited that bimodal bilinguals may require greater cognitive control to manage the switching between their two languages in different modalities (oral vs gestural) than same-modality (spoken–spoken) bilinguals. They also found that the GM volume in this area positively correlated with the degree of functional activation of left CN in response to a language switching task (switching between orally producing a picture name and signing the name). Pliatsikas et al. (2013) further examined past tense verb processing in Greek learners of English. Their learners performed a lexical decision task in which the target words (regular and irregular verbs) could be primed by a morphologically related word (the past tense form) or an unrelated word. The critical trials were target regular verbs primed by morphologically related words. The authors found that speed of processing on this critical trial type negatively correlated with GM density in the CB (and bilinguals showed more GM in the CB than monolinguals), and years of exposure correlated positively with GM density in the bilateral posterior putamen. In general, these data support a distinct role of proficiency or performance level in inducing structural brain changes, on top of the role played by L2 AoA (see also further evidence from training studies discussed below).

2.5. Structural brain changes induced by short-term L2 learning or training

Most of the studies reviewed above have examined bilingual speakers who have long-term (sometimes lifelong) experiences with a second language, either through immersion learning in an L2 environment or through extensive L2 use or L2 formal instruction. There are other individuals who acquire a second language on a relatively short-term basis, such as those who must take a job abroad or who have to migrate to a new country. Does short-term learning or intensive training lead to the same type of anatomical changes that we saw in

the previous studies? Several recent studies have been conducted to answer this question, and the evidence so far is encouraging.

Schlegel et al. (2012) studied American college students through the use of a longitudinal design to monitor differences in WM reorganization in adults in response to L2 learning. The participants underwent monthly DTI scans for nine months. Eleven participants signed up for a 9-month Modern Chinese intensive language course which met for 7.5 h a week. Results showed greater FA values for the students learning Chinese as opposed to the 16 control participants who did not learn Chinese within the same period. Importantly, there was a positive correlation between these Chinese learners' FA slope changes and the amount of Chinese they had learned, indicating that the structural changes were directly associated with language performance. The training group also showed lower RD scores, which the authors interpreted as indicating, when combined with the increased FA scores, an increase in myelination of these tracts. Further, five of their WM tracts that displayed learning-induced changes ended in the CN, in line with other studies implicating the CN in language and control functions (Grogan et al., 2009; Zou et al., 2012; see Sections 2.3 and 2.4).

Mårtensson et al. (2012) recruited a group of students (the "interpreters") from a military academy who went through an intensive language training program for 10 months in preparation for a career as military interpreters, and examined their behavioral performance and brain data across the first 3 months of training. The language training consisted of primarily the learning of vocabulary and idioms (roughly 300 to 500 weekly). Performance of the interpreters was compared with that of a control group matched for age and cognitive abilities (non-interpreter students). Over the course of 3 months, the interpreters, as compared to controls, showed increased CT in the left IFG, left middle frontal gyrus (MFG), and left STG, as well as increased right hippocampal volume. In addition, L2 proficiency in the trained language positively correlated with CT in the left STG and the right hippocampus (HP), indicating that sound-form mapping and new word formation might be distinctly important, given the emphasis of the training program on rapid vocabulary acquisition. In another study of interpreters, Elmer, Hänggi, and Jäncke (2014) found GM volume differences between multilingual controls and professional simultaneous interpreters who had varying number of years of interpreting experience (see also Elmer, Hänggi, Meyer, & Jäncke, 2011 for WM differences between the two groups). Somewhat counter-intuitively, the interpreters showed smaller GM volume and lower WM integrity than multilingual controls in a number of language and control brain regions and tracts. With respect to GM differences, Elmer et al. found that GM volume in the right IFG *pars opercularis* (IFGop), left IFG *pars triangularis* (IFGpt), middle ACC, and bilateral CN negatively correlated with cumulative number of years interpreting, suggesting that these differences might be a result of interpreting experience while other GM and WM differences were not.

The intensity with which the interpreters acquired a second language represents a somewhat extreme case of language learning, so it is unclear whether these findings can be easily generalized to other L2 learners who have more time and leisure to learn. In contrast to these classroom-based intensive learning programs, several other studies have trained participants in the lab, asking them to learn a simplified natural language (e.g., sounds or words) or sometimes an artificial language (e.g., Veroude, Norris, Shumskaya, Gullberg, & Indefrey, 2010; Wong, Perrachione, & Parrier, 2007; Yang & Li, 2012). The advantage of the training paradigm is that participants' learning process can be brought under tight control, so that they may be learning exactly the same material with exactly the same amount of time while their motivation and language background measured and accounted for (often with a control group for comparison). The training paradigm also presents distinct advantages for collecting brain data at precisely designated times, for example, scanning before training, midway into training, right after training, or with a time delay post training (see Fig. 2 for an illustration of the training paradigms used in such studies).

Hosoda, Tanaka, Nariai, Honda, and Hanakawa (2013) took advantage of the training paradigm, and brought native Japanese speakers who had begun to learn English at age 11 into the lab for a 16-week session to learn L2 English vocabulary. They included a control group that did not receive vocabulary training during the same period. After training, the learners, but not the control participants, showed increased GM and WM density in the right IFGop, and the increases correlated positively with their performance in L2 vocabulary. Interestingly, in a follow-up scan one year later, GM density returned to pre-training levels for those individuals who did not maintain language practice, but continued to increase for those who did practice their vocabulary. These results indicate that even in a laboratory setting, L2 instruction confers benefits in neural structure, and that continued use and practice is critical for the maintenance of these benefits.

In another lab training study, Wong et al. (2008) trained adult monolingual English speakers to learn pseudowords with pictures. Key to this study was that the pseudowords were associated with different pitches that conveyed meaning, such as what happens for lexical tones in Chinese. Participants received training over the course of weeks until they reached high-proficiency in the task. Based on how they reached the final learning criterion (e.g., greater than 95% performance), participants were classified as "successful" and "less successful" learners, respectively. The successful learners, as compared with the less successful learners, showed larger GM volume, and a trend towards more WM, in the left Heschl's Gyrus (HG), part of primary auditory cortex. Moreover, performance on the word-pitch mapping task correlated positively with GM and WM volume in the left HG. Importantly, these data matched with the results from a functional MRI study by Wong et al. (2007), in which successful learners showed more focused activation in the left STG, an area crucial for phonological processing, whereas the less successful learners were characterized by a more diffuse network in the frontal and temporal cortical regions. The two groups also showed differences in activation patterns even before learning (see further discussion of individual differences in Section 3.5).

The studies discussed above revealed significant anatomical changes in terms of increased GM density, increased CT, or enhanced WM tract connectivity, not as a result of lifelong
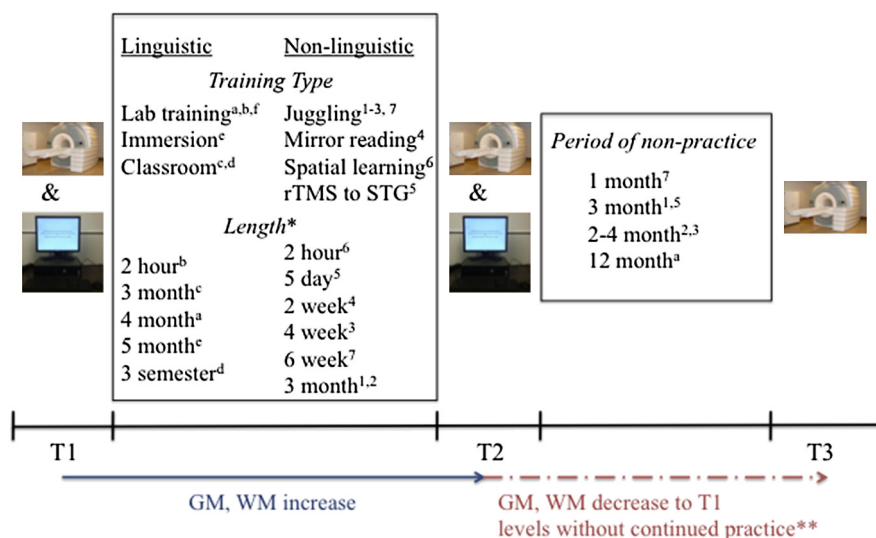
Fig. 2 – Overview of protocols used in the language training studies ([a]Hosoda et al., 2013; [b]Kwok et al., 2011; [c]Mårtensson et al., 2012; [d]Schlegel et al., 2012; [e]Stein et al., 2012; [f]Wong et al., 2008) and non-linguistic training studies ([1]Boyke et al., 2008; [2]Draganski et al., 2004; [3]Driemeyer et al., 2008; [4]Ilg et al., 2008; [5]May et al. (2007); [6]Sagi et al., 2012; [7]Scholz et al., 2009). At time points 1 (T1) and 2 (T2), participants undergo both behavioral testing and structural scanning, and time point 3 (T3), participants are scanned only. [*]Training lasted a variable amount of time in Wong et al., 2008; scans were performed monthly in Schlegel et al., 2012. [**]GM increase in Scholz et al., 2009 found between T1 and T2 continued to T3 without juggling practice. See Appendix for abbreviation keys to the brain regions.

experience with two languages but as a function of short-term learning or training on aspects of a new language. Thus, short-term learning can indeed modify brain structures. One important question in this regard, rarely asked in the neural studies of bilingualism, is whether the context of learning will significantly impact the brain structure as well as the performance on a second language. The traditional learning context such as the classroom may be less conducive of an environment than the realistic language context: in the latter case, the learner is immersed, has direct access to native speakers, and has richer and more multi-modal interactions with the L2 environment and the native speakers of the target language. To examine L2 learning in an immersion context, Stein et al. (2012) tested college students who went to Switzerland for study abroad experience and who learned German over the course of 5 months. Their results indicated that at the end of their study abroad experience, the learners had increased GM density in the left IFG and the left anterior temporal lobe (ATL), two areas that are implicated in lexical access and semantic integration. The structural changes in these regions also positively correlated with the students' performance on L2 vocabulary tasks. However, this study lacked a control group, which weakens the claim that the brain changes are due particularly to L2 learning in the immersion context and not other potentially confounding factors (e.g., mixed L1 language background or motivation or anxiety associated with living in a new country).

3. Mechanisms of anatomical change

The studies reviewed so far provide convincing evidence that neuroanatomical changes can be induced by experiences with

learning a new language. What remain to be understood are the questions of when and how such changes may occur, what learner variables modulate these changes, and what environmental factors may enhance, attenuate, or else minimize such changes. In the above review we have already touched on some of these "mechanisms of change" questions; in this section we specifically raise these questions in order to provide a synthesis of the various studies, and to build a basis for future research directions in this domain. We also make an attempt to connect and compare the studies reviewed above with non-linguistic studies of other cognitive skills, although a full review of non-linguistic studies is beyond the scope of this article (but see reviews in Lövdén, Wenger, Mårtensson, Lindenberger, & Bäckman, 2013; May, 2011; Thomas & Baker, 2013; Zatorre et al., 2012).

3.1. Do structural brain changes mirror functional neural activity patterns?

Because the study of structural brain changes induced by L2 learning is a relatively new enterprise, only a few studies have directly compared functional neural patterns with structural brain data within the same experiments (e.g., Abutalebi et al., 2012; Gold et al., 2013; Hosoda et al., 2013; Luk et al., 2011; Wong et al., 2007, 2008). Even with the limited evidence, we can start to ask the question of whether anatomical changes in the brain are mirrored by functional neural patterns, or whether our existing knowledge of functional brain activity is consistent with the observed anatomical data.

Given that expanded GM density is the most commonly identified anatomical change in the data reviewed so far, in Fig. 1, we present a summary of the common areas where

increases in GM have been found. As the figure indicates, bilingual individuals who have had significant experience learning an L2, whether through natural setting, formal instruction, or lab training, in general show GM volume/density increases, in several key areas that have been previously reported in functional neuroimaging studies, including the IFG, MFG, STG, ATL, IPL, all in the LH, the CB and the HP in the right hemisphere (RH), and the CN, mostly in the LH but in some cases bilaterally (see also Table 2).

Almost all of these areas have been shown to be crucial for language learning or processing in the functional neuroimaging literature. Below is a brief summary of their functions based on a number of reviews or meta-analyses of functional neuroimaging data (mostly from fMRI studies), in particular, Binder and Desai (2011), Price (2000; 2010), Hickok (2009), Hickok and Poeppel (2007), and Rodriguez-Fornells et al. (2009). (See Appendix for a list of abbreviations of the brain regions.)

IFG: a core brain region for language in the LH, involved in lexical retrieval, articulatory planning, and morpho-syntactic processing; further subdivisions within IFG may be recruited for different processes of bilingual language production (Parker Jones et al., 2011);

MFG: this area in the LH may sometimes overlap with the dorsolateral prefrontal regions (e.g., DLPFC), which may be important for word meaning selection, along with functions for articulatory planning and executive control (e.g., working memory and response inhibition);

STG: mainly involved in acoustic and phonetic/phonological processing (e.g., tone processing in Chinese), with the more dorsal part being more sensitive to basic acoustic characteristics while the more posterior part important for phonological encoding (Zhang et al., 2011);

ATL: this area in the LH, along with both inferior and middle temporal gyrus (ITG, MTG), has been shown to handle various aspects of lexical semantic representation and processing;

IPL: this region in the LH, including adjacent areas encompassing the SMG and AG, have been implicated in phonological working memory, phonological storage, semantic integration, and thus vocabulary learning in general; the AG might be particularly involved in event representation and episodic memory retrieval (Binder & Desai, 2011);

HP: critical for memory formation and hence formation of new lexical−semantic associations;

CN: as part of the BG structure, the CN plays an important role in sequence learning, procedural memory (e.g., Ullman, 2001, 2004), phonological learning, and detection of phonological anomalies (Grogan et al., 2009);

CB: although not a classic language area, cerebellum has been recently shown to be involved in lexical semantic processing (Li, Jin, & Tan, 2004; Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995), and in grammatical processing as part of the procedural memory system.

Given the role of these regions for the neurobiology of language, the findings that these are the areas showing increased GM as a function of L2 experience are not surprising. In addition, previous work in the functional neuroimaging of

bilingualism has also suggested a network of cortical and subcortical structures in the LH that are engaged in bilingual language monitoring and control, specifically a frontal-striatal circuit involving IFG, the ACC, the IPL, and the CN. As discussed in Section 2.3, recent structural imaging data also confirm that this functional network of bilingual control has its concomitant anatomical substrates, with enhanced GM density in the ACC, IPL, and CN as a result of bilingual learning and experience (e.g., Abutalebi et al., 2012; Della Rosa et al., 2013; Grogan et al., 2009; Zou et al., 2012). These data allow us to conduct three-way comparisons and correlations among functional brain activity, anatomical brain change, and behavioral performance, so that we can identify, significantly more likely, direct relationships between experiences and adaptive brain structures. Finally, there is also evidence that structural connectivity based on DTI data and functional connectivity patterns from fMRI data are compatible (see discussion below). In general, we found a highly consistent picture for structure and function correspondence, suggesting that structural neuroplasticity is a result of the dynamic functional engagement and adaptation (see also this correspondence in non-linguistic studies reviewed in Lövdén et al., 2013; see also Section 3.2).

A related question to ask is whether there are also correspondences within the structural data as revealed by GM volume, WM integrity, and CT. Most of the reported studies have so far focused on one of the three types of methods, and only in a few cases (e.g., Gold et al., 2013; Golestani, Molko, Dehaene, LeBihan, & Pallier, 2007; Hosoda et al., 2013), researchers have used a combination of measures to address whether changes in GM density, CT, and WM integrity co-occur or co-localize (though no study has looked at all three). For example, as mentioned in Section 2.2, Gold et al. found significant differences only in WM integrity but not in GM volume between bilinguals and monolinguals, suggesting that GM and WM changes may occur independently. Golestani et al. (2007) showed that increased GM density in a region anterior to the left parieto-occipital sulcus (POS) was accompanied by decreased WM density in that region (see further discussion in Section 3.5). A high GM-WM correspondence was found in Hosoda et al. (2013; see Section 2.5): Japanese learners of English had increased GM volumes in a number of regions including the IFG, CN, and STG/SMG after training, and importantly, the GM changes corresponded with the patterns of WM connectivity between the right IFG with CN and with STG/SMG. In general, given the small number of multiple structural imaging methods within the same study, there appear to be few significant correspondences across multiple measures of structural changes (although encouraging data are emerging in non-linguistic studies, e.g., Scholz, Klein, Behrens, & Johansen-Berg, 2009; see Section 3.2).

The study of structure-function correspondences also points to the need to identify a more direct causal link between specific second language experience and neuroplasticity. Functional neural activation patterns tend to be mostly correlational when evaluated against learning variables (e.g., increase *vs* decrease in left IFG may mean either better or worse L2 performance; see Li, 2014 for a discussion), but anatomical structure changes in the brain, such as GM and WM density,

may be more easily interpreted in causal terms when they can be correlated with performance/proficiency levels: for example, decrease versus increase in GM has more distinct meaning than decrease versus increase in Blood-Oxygen-Level Dependent (BOLD) signal (but see Kanai & Rees, 2011; Box 3, for a cautionary note in this regard). Such structural changes can be especially informative when longitudinal studies are designed, through a training paradigm, to track brain changes within the same individual over a specific period of time (e.g., in Della Rosa et al., 2013), and to compare groups with versus without the experience of interest (i.e., training *vs* control groups; e.g., in Mårtensson et al., 2012). Thus, future studies should conduct more longitudinal structural imaging work to establish more direct, hopefully causal, relationships between bilingual experience and neuroanatomical change (see Li & Green, 2007 who called for attention to the important role of longitudinal neuroimaging research in bilingualism).

3.2. Is language experience unique in inducing structural brain changes?

Language learning can be an intensive experience occurring on a daily basis and across the lifespan, and as such, it provides a powerful environmental input to the nervous system to induce anatomical changes in the human brain. The question we need to ask is whether language experience is unique in bringing about structural brain changes, or whether language, as part of the cognitive system (in contrast to the proposal of Fodor, 1983), works in the same way as other components of cognition that also give rise to structural changes in the brain. Many previous studies have identified anatomical changes as a function of non-linguistic experiences or acquisition of new skills, including attention, musical expertise, mathematical learning, spatial memory, and visuo-motor learning. It is not the goal of the current paper to review the large number of non-linguistic studies, as several recent papers have already attempted to provide integrative reviews of experience-dependent structural changes in the brain (Lövdén et al., 2013; May, 2011; Thomas & Baker, 2013; Zatorre et al., 2012). Here we present a summary of the studies that have used a training paradigm, primarily with examples from spatial memory and visuo-motor training, in order to identify common mechanisms underlying experience-dependent structural brain changes in both linguistic and non-linguistic domains. Fig. 3 presents an overview of the brain regions affected by training in studies of linguistic as compared to the non-linguistic experiences.

An earlier well-known imaging study of structural neuroplasticity is Maguire et al. (2000), in which London taxi drivers' navigation and spatial memory capacities were examined. Maguire et al. showed that taxi drivers, compared with age-matched non-taxi drivers, had increased GM volume in the right posterior HP, and the size of this increase correlated positively with the taxi driver's time spent driving. These structural brain data based on GM also matched well with the same authors' functional imaging data, in which the right posterior HP was recruited more strongly in successful navigation trials than in non-successful trials (in a virtual reality experiment). The authors suggested that the HP volume changes reflected the taxi drivers' experience with their detailed memory representation of the city, and that the right posterior HP could be the storehouse of spatial representation of the environment.

To establish a more substantiated relationship between experience and neural change, Maguire and colleagues conducted two additional studies, examining GM volume in non-taxi drivers (Maguire et al., 2003) and bus drivers (Maguire, Woollett, & Spiers, 2006). For the non-taxi drivers, unlike the taxi drivers, no differences in GM volume were found to be associated with inherent navigation expertise (non-taxi drivers who had varying levels of navigation skills), suggesting that increases in HP volume are not related to inherent or innate navigation ability per se. For the bus drivers, whose driving is more spatially confined (i.e., route following), there was also no association between GM volume difference and the amount of time (in years) driving. This suggests that it is not the driving experience per se, but rather flexible learning, the representation and use of spatial knowledge as in taxi driving, that leads to enhanced spatial memory and consequently brain changes. In short, these studies together portray a picture of how specific navigation experience affects spatial memory that results in neural changes.

A number of other studies are also consistent with the picture of Maguire and colleagues' studies. For example, Draganski et al. (2004) trained a group of participants to learn the 3-ball cascade juggling routine for a 3-month period, and performed brain scans before and after the training period. They compared the brain structures of the participants (the jugglers) and a control group (the non-jugglers), and found that after training the jugglers had a significant bilateral expansion in GM density in the mid-temporal area and in the left posterior intra-parietal sulcus, whereas the non-jugglers showed no such changes. In two other similar longitudinal studies, Driemeyer, Boyke, Gaser, Büchel, and May (2008) and Scholz et al. (2009) trained participants on the same juggling routines for 4 and 6 weeks, respectively, and also found that the jugglers had significant learning-related increases in GM density or in WM tracts overlapping the same visual and motor integration areas. Scholz et al. (2009) additionally showed that the increases in GM density co-localized with the cortical regions where WM tracts had significant FA increases, although there was no correlation between the magnitude of GM and WM changes.

Many other studies of selective attention, vision, mathematical learning, musical expertise, and literacy acquisition have also shown structural neuroplasticity in the adult brain as a function of cognitive task or skill acquisition. To illustrate, Schneider et al. (2002) found significantly more GM volume in the anterior portion of the HG in musicians (professional and amateur) than in non-musicians. Carreiras et al. (2009) showed that adults who acquired literacy late in life (at a mean age of 32 years), as compared with age-matched illiterate adults, had significantly more WM in the splenium of the CC, and increased GM volume in a number of regions implicated for reading skills. Furthermore, these authors found a close correspondence between WM and functional connectivity patterns, indicating that reading, even when acquired late in adulthood, can enhance inter-hemispheric connections especially in the left and right AG. These studies from musicality and literacy, among many others, provide additional data consistent with the findings from spatial memory and visuo-motor skill
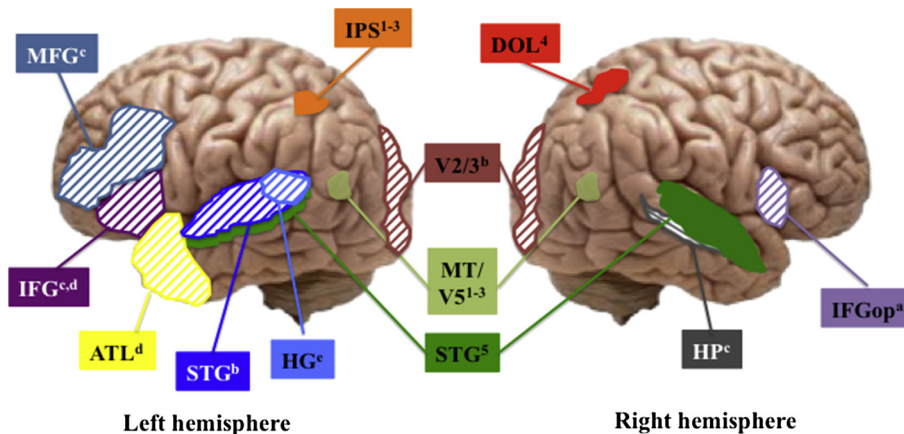
Fig. 3 – Representation of brain regions that showed GM increase as a result of linguistic (shaded) or non-linguistic (solid) training. Linguistic studies include [a]Hosoda et al., 2013, [b]Kwok et al., 2011; [c]Mårtensson et al., 2012, [d]Stein et al., 2012, and [e]Wong et al., 2008; Non-linguistic studies include [1]Boyke et al., 2008, [2]Draganski et al., 2004, [3]Dreimeyer et al., 2008, [4]Ilg et al., 2008, and [5]May et al., 2007. * The hippocampus is a medial structure, but it is identified here laterally for illustrative purposes. See Appendix for abbreviation keys to the brain regions.

acquisition discussed above (see further reviews in Bavelier, Green, Pouget, & Schrater, 2012; May, 2011; Zatorre et al., 2012; but see Lövdén et al., 2013; Thomas & Baker, 2013 for cautionary notes and methodological considerations). In light of these data, we can conclude that language experience is not unique in inducing structural brain changes, and that structural neuroplasticity can occur as long as the relevant experience provides sufficient stimulation to the brain, regardless of whether the simulation is from language or non-language tasks.

3.3. Is long-term, sustained, experience required for structural brain changes?

Our discussion above suggests that both linguistic and non-linguistic experiences serve to modify structural properties of the brain. But how much experience is needed for anatomical, as opposed to functional, changes in the brain to occur? When do we observe long-term, sustained, effects and when do we see short-term, transient, effects? How do these long-term versus short-term effects vary as a function of type, amount, intensity, and time of experience?

In the L2 studies discussed in Section 2, most training or longitudinal experiments lasted for a few months to a year, and the shortest term within which structural brain effect has been observed is 3 months. In non-linguistic training studies, the timescale can be much shorter. For example, in the Draganski et al. (2004) and Scholz et al. (2009) studies discussed above, both GM and WM changes were observed after 3 months and 6 weeks of juggling training, respectively. In the Driemeyer et al. (2008) study, clear transient effects in structural brain changes were observed as early as after 7 days of training of the juggling task (with GM expansion in the occipito-temporal cortex). Could there be a difference between language experience and other cognitive skills experience in terms of the time needed for structural brain changes to occur?

The non-linguistic studies certainly seem to indicate that experience-dependent changes may occur rapidly and transiently in the brain, for example, appearing after only a few weeks of learning. Effects from even shorter timeframe have been observed: training regimens as brief as 15 min per day for 2 weeks have been shown to induce GM volume changes. For example, Ilg et al. (2008) showed that after only 2 weeks of training to read mirrored words projected on slides, participants had increased GM in the right dorsolateral occipital cortex during mirror reading. This also corresponded to the site of the peak activation from their fMRI data during the mirror-reading task. In another study, May et al. (2007) showed that repetitive transcranial magnetic stimulation (rTMS) to the superior temporal cortex could cause GM to increase in the auditory cortex within only 5 days of continued presentation. A recent study reported an effect of training with perhaps the shortest timescale: Sagi et al. (2012) showed that brain changes, as measured by DTI in terms of decreased MD and increased FA, could occur in the left HP and bilateral parahippocampus with only 2 h of training of participants on a spatial learning and memory task (e.g., in a computerized car racing game).

Thus, the non-linguistic data so far suggest that structural neuroplasticity could occur much more quickly and transiently than we have previously thought. One could argue that complex visuo-motor skills such as juggling or mirror reading are highly cognitively demanding, and therefore provide a stronger stimulation to the visual and motion-sensitive cortical regions and induce structural brain changes more easily than do language tasks in short-term. The fact that language training usually takes longer to see an effect (e.g., 3-month intensive training as in Mårtensson et al., 2012) could suggest a difference between linguistic and non-linguistic tasks. However, a recent study by Kwok et al. (2011) showed very rapid changes in GM volume resulting from intensive learning of four color names: participants were taught four "new words" (monosyllables) that were each associated with one of four different color shades (two types of blue shades

and two types of green shades). Their results indicated that after 5 sessions of learning, each lasting an average of 20 min (less than 2 h in total), GM volume increased in the V 2/3 of the left visual cortex, a region known to mediate color vision. Although the participants were learning only 4 new words, which is by no means comparable to any real language learning task, this is perhaps the most rapidly produced structural change reported so far in the language domain.

Comparisons of linguistic and non-linguistic studies also raise the question of whether the timescale of effects is directly tied to the timescale of experience. For example, do short-term training yield only transient effects, and the shorter the training, the more quickly the effect attenuates? Do long-term experience (such as years of language learning) have more long-lasting effects? Driemeyer et al. (2008) suggested that the exact timescale of experience-dependent brain changes will depend on the intensity and quality of the experience of the learner: brief experience can have an effect only when the learning/training experience is intensive and involves qualitative changes in behavior (see further discussion in Sections 3.4 and 4 on linguistic *vs* non-linguistic differences in complexity, intensity, and task demand). The evidence so far in this regard for non-linguistic training remains limited and mixed: Driemeyer et al. (2008) and Draganski et al. (2004) showed that the jugglers had decreased GM volume after an extended period of no juggling practice (2—4 months), whereas in Scholz et al. (2009), after 4 weeks without juggling, the jugglers still had highly elevated GM volume. A related question to ask, in the case of linguistic tasks, is whether different contexts of learning (e.g., immersion learning, virtual environment, traditional classroom) will give rise to different outcomes of learning and memory, and the corresponding functional and structural brain changes (see discussion of Stein et al.'s 2012 study in Section 2.5, and also discussion in Section 3.5). Perhaps more embodied experiences such as those through immersion and virtual environments will lead to more long-lasting brain effects as well as better learning outcomes. Such questions need to be addressed in future studies.

3.4. Are age and proficiency of the learner both crucial for structural brain changes?

In most of the studies reviewed so far, the participants are typically young adults in the age range of 18—30 (the majority being college students). While evidence points to clear malleability of brain structures at this age range, and while we are certain that the brain is highly adaptive before this age, what happens after this age?

A few studies have recently examined anatomical changes in the aging brain as a function of L2 training or learning experience. In Section 2.2 we reviewed the study of Abutalebi et al. (2014), Luk et al. (2011), and Gold et al. (2013) regarding GM volume and WM integrity in the elderly, which suggests that bilingual experience might provide a protective mechanism for age-related cognitive decline. These studies, however, do not directly address the question of whether short-term learning in older adults can result in structural brain changes as in young adults, given that their bilingual participants had lifelong experience with two or more languages. Boyke, Driemeyer, Gaser, Büchel, and May (2008)

addressed this question head-on. Older participants (mean age, 60 years) in their study were trained to learn the juggling task in 3 months, as in Draganski et al. (2004). Surprisingly, GM volume changes were similarly observed in the mid-temporal areas of the visual cortex as in young adults (mean age, 20 years) in the Draganski et al. study. In addition, the older adults showed GM increases in the HP and the BG, two important areas for new skills acquisition, particular for memory formation and sequence and motor learning. Although the older participants showed a smaller GM effect when directly compared with younger adults, these data in general suggest that age per se is not limiting factor for structural neuroplasticity. These are highly encouraging data for aging individuals who consider to learn a new skill. In light of such findings of neuroplasticity in older adults, Antoniou et al. (2013) proposed that foreign language training for older adults could serve as a vehicle to strengthen neural networks and potentially mitigate age-related cognitive declines, as this type of training may engage a larger neural network than activities such as math or crossword puzzles. This is consistent with the idea discussed earlier that neural changes of both WM and GM in key language control areas may provide a protective mechanism, a "cognitive reserve" as a deterrent to neurodegenerative diseases (Alladi et al., 2013; Bialystok, Craik, & Freedman, 2007; Gold et al., 2013; Luk et al., 2011).[3]

Parallel to the age effect in the language domain is the effect of proficiency in a new language. In the functional imaging literature, there have been intense debates regarding the relative contribution of AoA versus proficiency in modulating bilingual language activation patterns (see Hernandez, 2013; Hernandez & Li, 2007; Li, 2014 for reviews). Earlier studies suggested that AoA is a more important variable (Kim et al., 1997; Mechelli et al., 2004; Weber-Fox & Neville, 1996), whereas later studies point to the distinct role of proficiency (Abutalebi, 2008; Abutalebi et al, 2014; Abutalebi & Green, 2007). As discussed in Section 2.5, the structural imaging data also showed a correlation between the learner's level of L2 proficiency or other behavioral measures of performance and the extent of GM and WM changes (e.g., Grogan et al., 2012; Mårtensson et al., 2012; Mechelli et al., 2004; Pliatsikas et al., 2013; Schlegel et al., 2012; Wong et al., 2008; Zou et al., 2012). In addition, performance levels on linguistic tasks are often correlated with structural brain variability in the individual learners (see Section 3.5 for further discussion). Overall, based on the extant literature we can probably conclude safely that both AoA and proficiency are crucial variables for language learning, and that the two variables may interact in intricate ways or work differently for different domains (see Hernandez & Li, 2007; Mechelli et al., 2004; Wartenburger et al., 2003). Behind the effects of both AoA and proficiency must be the story of neural efficiency, due to increased or decreased engagement of the language network and the attentional control resources for obtaining the same processing outcome (see Hernandez, 2013, Chapter 4; for a discussion).

---

[3] Even simple exercises such aerobics (e.g., for six months) can improve memory and increase the size of hippocampus, as shown in a recent study with a group of older participants (aged 55—80; Erickson et al., 2011). See Lövdén et al. (2013) for a review of several training studies in the aging population.

In non-linguistic domains, the relative role of age versus proficiency or performance level is so far unclear. In particular, researchers found mixed results with regard to the correlation between performance level and magnitude of GM/WM changes. In Schneider et al. (2002) and Aydin et al. (2007), GM (but not WM) volume was positively and significantly correlated with standardized musical aptitude and with years of being mathematicians, respectively. In Maguire et al. (2000), GM volume was found to correlate with years of taxi driving experience and in Roberts, Bain, Day, and Husain (2013), WM integrity with both amount and AoA of karate experience for karate experts. But in several other studies including Draganski et al. (2004), Boyke et al. (2008), and Scholz et al. (2009), no GM or WM change was found to be correlated with performance level or training progress of juggling (see also reviews in Lövdén et al., 2013; Zatorre et al., 2012). These differences raise the question of whether linguistic and non-linguistic performances differ in bringing about structural brain changes, but more importantly the question of what performance level is required for anatomical changes to take place in the brain.

In Section 3.3 we discussed the question of short-term versus long-term effects, and indicated that experiences as short as hours and weeks, not months and years, could produce significant brain changes in non-linguistic domains. Is it possible that there is a better correspondence between performance/proficiency and functional and structural brain changes for language, simply because one needs more linguistic experience to gain some proficiency in the L2 and therefore more time for functional and structural brain changes to occur? This seems likely, given the complexity of the human language system, and the corresponding neurocognitive demands in learning a new language. The fact that language tasks can engage a widespread network in frontal, temporal, parietal, and striatal regions speaks to the complexity as well as the cognitive and memory demands that language places on the learner. The study of the similarities and differences between language and non-language tasks in impacting the brain will also have significant implications for understanding the so-called 'temporal parameter': the amount of experience that is needed for real anatomical changes to occur in the brain, whether linguistic or non-linguistic. Researchers are actively pursuing this question in several non-linguistic domains (e.g., Draganski et al., 2004; Driemeyer et al., 2008).

To fully understand the temporal parameter issue, we will also need to identify the cellular and molecular mechanisms underlying experience-dependent structural brain changes (see discussions in Antoniou et al., 2013; Kanai & Rees, 2011; May, 2011; Zatorre et al., 2012), as it is certain that structural brain changes differ as a function of age. It may also be more prone to certain aspects of learning or task demands (e.g., spatial memory and navigation), due to the microstructural properties of the target brain regions (e.g., HP, in which hippocampal neurogenesis is rampant throughout human adulthood; Spalding et al., 2013). Currently, the neurobiological basis of macroscopic structural changes is not well understood, and what GM and WM differences entail at a microscopic level remain unclear: for example, as mentioned in 2.1.1, GM increases could mean increases in cell size, neuronal or glial cell genesis, angiogenesis, synaptogenesis, or even changes due to increased cerebral blood flow. Thus, the aggregated measurement of GM, although useful, is not highly informative.

3.5. Can structural brain changes capture individual differences in L2 learning?

A number of functional MRI studies have already investigated the issue of individual differences in L2 learning (e.g., Sheppard, Wang, & Wong, 2012; Ventura-Campos et al., 2013; Veroude et al. 2010; Wong et al., 2007; Yang & Li, 2012; Yang, Gates, Molenaar, & Li, 2014). These studies indicate that functional neural patterns observed in L2 learners can indeed capture individual differences, and in some cases, even predict learning successes before L2 experiences begin. They also suggest that the individual differences in language learning might be associated with different cognitive abilities, including working memory and cognitive control as reflected in differential activations of the brain's executive network (see Section 2.3). In terms of methods, researchers have used a variety of whole-brain and ROI-based analyses, resting-state functional connectivity analyses, and brain network (e.g., small-world networks) analyses, and these analyses are conducted within a pre- versus post-training paradigm (see Li, 2014 for a review). Such methods allow researchers to differentiate the good/successful learners from the poor/less successful learners by comparing pre- versus post-training functional neural patterns.

Compared with the increasing number of functional MRI studies in addressing individual differences, relatively fewer structural MRI studies have been conducted to examine the individual differences in GM, WM, or CT changes as a function of L2 learning. Table 3 presents a summary of existing studies that have specifically addressed the issue of individual differences from the structural neuroplasticity perspective. Golestani, Paus, and Zatorre (2002) asked Spanish-speaking participants to learn a non-native speech contrast (Hindi dental-retroflex contrast), and found that a faster rate of learning was related to greater WM in parietal regions, especially in the LH. The same patterns were found in Golestani et al. (2007): faster learners, as compared to slow learners, showed a number of structural brain differences, and in particular higher WM density and larger WM volume in the left HG. Golestani and Pallier (2007) also examined whether there were GM or WM differences associated with the accurate pronunciation of non-native phonemes, and identified increased WM (but not GM) density in the left insula/prefrontal cortex and bilateral IPL regions for accurate as compared with non-accurate speakers.

The role of HG as reported for learning phonetic contrasts is highly meaningful in light of the findings from Wong et al. (2008), where successful learners of lexical tones, compared with less successful learners, showed greater GM volume and a trend towards increased WM volume in the left HG (see discussion in Section 2.5). However, Sebastián-Gallés et al. (2012) found that when bilingual Spanish-Catalan speakers were divided into the good performers and the poor performers in the perception of L2 phonetic contrasts, they had no GM differences in the HG. Sutherland et al. (2012) further investigated whether structural differences in HG may be related to the detection of frequency modulation in tones, an ability

Table 3 – Individual difference studies of structural brain variation related to language learning and processing.

Study	Behavioral task	Relationship to Structure[a]
Phonological		
Golestani et al., 2002	Learning to perceive non-native phoneme contrast	(+) WM parietal lobe, especially left hemisphere
Golestani et al., 2007	Learning to perceive non-native phoneme contrast	Faster > Slower learners: WM density, volume in HG
Sebastían-Gallés et al., 2012	Perceiving non-native phoneme contrast	Poor > Good learners: WM right insulo/fronto-opercular region; No GM
Golestani & Pallier, 2007	Pronunciation of non-native phonemes	Good > Poor pronouncers: WM left insula/prefrontal cortex, bilateral IPL; No GM
Wong et al., 2008	Learning lexical tones for word identification	More > Less successful learners: WM volume HG
Sutherland et al., 2012	Detection of frequency modulation of tones	(+) GM HG for 10-year-olds and 13-year old boys
Grammatical		
Xiang et al., 2012	Grammatical inferencing	(+) FA in BA45-posterior temporal lobe pathway
Flöel et al., 2009	Artificial grammar learning	(+) FA in tracts originating from Broca's area
Loui et al., 2011	Artificial phrase-structure rule learning of tones	(+) Tract volume ventral AF Predicted by WM in right SMG
Fluency		
Grogan et al., 2009	L2 Phonemic fluency	(+) GM bilateral head of CN
	L1 and L2 Semantic fluency	(+) GM ITC

See Appendix for abbreviation keys to brain regions.
[a] Correlations are denoted with '(+)' for positive relationships and '(−)' for negative relationships.

underlying speech processing. A unique feature of this study is their use of a longitudinal design to track children at three time points, at 10, 11.5 and 13 years of age. Their data indicated that an individual's performance in frequency modulation detection correlated with GM density in the left HG for the youngest age (age 10), but this correlation weakened with age.

Overall, the above studies suggested that pre-existing structural variability in the left HG may serve as a powerful predictor of phonetic learning and phonological processing abilities in children and adults, given the importance of these abilities for auditory perception, speech comprehension, and production. However, there remains the possibility that the structural variability itself may be driven by environmental stimulation provided by language. For example, Ressel et al. (2012) compared early simultaneous Spanish-Catalan bilinguals to Spanish monolinguals to see whether HG volume

might be modified by the effect of early L2 experience. Their VBM analysis revealed larger GM volumes in bilinguals than in monolinguals. Since the two groups were matched on education, socio-economic status, and musical experience, this HG volume difference was taken to indicate that second language learning experience played a causal role in the increased size of the auditory cortex, rather than the other way around (i.e., different size of HG affecting learning success).

While most of the structural brain studies of individual differences have focused on how structural variability affects phonological learning, Xiang et al. (2012) examined individual differences in a variety of linguistic domains as measured by standardized language aptitude tests that included sound recognition, vocabulary learning, grammatical processing, working memory, and IQ. Whether general language aptitude might be related to one's success in learning a second language has been an issue of interest to many applied linguists as well as cognitive scientists (see Dörnyei, 2005; Miyake & Friedman, 1998). Xiang et al. correlated individual performances on these tests with DTI measures (e.g., FA, structural laterality of connections). Their analyses revealed significant correlations between individual's performance and a number of structural pathways, specifically between the IFG regions (BA6, 44, 45, 47) and the temporal and parietal lobes; for example, the BA45 to posterior temporal lobe pathway significantly captured individuals' grammatical ability, while the BA47 to the parietal lobe pathway reflected rapid vocabulary learning ability.

In another study, Flöel, de Vries, Scholz, Breitenstein, and Johansen-Berg (2009) asked participants to learn a finite state grammar in which letter sequences of bigrams or trigrams were visually presented. They showed that learners' performance on grammaticality judgments of rule-based sequences correlated with the mean FA values in tracts originating from the Broca's area. Unlike Flöel et al. who used visual stimuli, Loui, Li, and Schlaug (2011) asked participants to listen to sequences of auditorily presented tones whose frequencies were governed by artificial phrase-structure rules. The participants were then tested on recognition of learned sequences and generalization to novel sequences. Their performance on the generalization, but not the recognition, correlated with tract volume in the ventral arcuate fasciculus that connects the right IFG and MTG. Moreover, this performance was predicted by FA value in the WM underlying the right, but not left, SMG. Overall, these studies are consistent with functional imaging studies that have also used artificial grammar learning paradigms in which individual differences were identified (e.g., Yang & Li, 2012).

In a review of the recent literature on neuroanatomical correlates of individual differences, Kanai and Rees (2011) examined evidence from a large number of studies of motor behavior, perception, intelligence, and personality. They suggest that investigation of individual differences in brain structure is a critical tool for understanding cognition and behavior, which we believe is also true in the domain of language. Future research should also examine the interaction of pre-existing structural variability and the effects of environmental input (amount, type, and nature of input; see Section 4) so as to arrive at a deeper understanding of why and how second language learning may be successful or not.

4. General discussion and future directions

Neuroplasticity in humans is characterized by the extraordinary ability of the human brain to adapt in response to environmental stimulus, cognitive demand, or behavioral experience. Cognitive brain research in the last decades, especially the past 20 years, has revealed exciting new findings regarding neuroplasticity at many different levels, across many different domains. The study of neuroplasticity as a function of the individual's linguistic experience has also led to important findings. The understanding of how adaptive neural changes are triggered by experiences with a new language and the interaction between the new and old languages, however, has only begun recently. In this article we have reviewed emerging evidence regarding how structural neuroplasticity occurs as a result of one's bilingual experience. The evidence reviewed so far shows that our linguistic brain is much more plastic than we have ever previously imagined, and the extant data from structural imaging studies portray a picture that is highly consistent with structural neuroplasticity observed in non-linguistic domains. Second language experience-induced brain changes, including increased GM density and WM integrity, can be found in children, young adults, and the elderly; can occur rapidly with short-term language learning or training; and are sensitive to age, age of acquisition, proficiency or performance level, language-specific characteristics, individual differences, and possibly other environmental and learning properties. At the same time, these patterns are also highly consistent with findings from functional imaging studies of L2 learning and bilingualism. Our review identifies the mechanisms that drive L2 experience-dependent anatomical changes, and integrates structural imaging evidence with our current knowledge of functional neural patterns of language and cognitive skills.

In a seminal paper that predated much of our current knowledge about structural neuroplasticity of language and bilingualism, Bates (1999) proposed a dynamic emergentist perspective for us to consider issues surrounding experience-dependent neuroplasticity, brain organization and reorganization, and language development. In the larger context of language acquisition, she described the dynamic interactions that occur between the developmental processes of neural structure, neurogenesis/synaptogenesis, brain maturation, cognitive and linguistic processes of learning-induced neural changes, and the implications that these interactions have for understanding the time course of both neuroplasticity and language development, especially with regard to the so-called "critical period" of language learning. Furthermore, Bates (1999) embraced "pluripotentiality", the idea that cortical tissues have the capability of taking on a wide range of representations, but the success of their doing so depends on the *timing*, *nature*, and *extent* of the input to which the tissues are exposed (see also Elman et al., 1996). This theoretical perspective, and its view of the three dimensions of input, are highly informative to the understanding of structural neuroplasticity induced by second language experience, given the evidence reviewed in this article.

First, the *timing* of second language learning is important, which we doubt anyone would dispute. The extant data from

both functional and anatomical brain studies suggest that L2 age of acquisition is an important factor, albeit not the "critical" factor as suggested by the "critical period hypothesis". The encouraging evidence for learning, whether language or not, is that the brain can continually modify and reconfigure its function and structure, even at a later stage, as reflected in changes in GM, WM, and connectivity among regions, as clearly indicated in our review above.

In recent years there have been alternative accounts of the critical or sensitive periods of language learning (e.g., Elman, 1993; Johnson & Newport, 1989). One stronger contender to the original biologically based account (Lenneberg, 1967) is the "competition model" (Bates & MacWhinney, 1987; Hernandez, Li, & MacWhinney, 2005; Li & MacWhinney, 2013; MacWhinney, 2012). In its various formulations including the computationally instantiated DevLex (Li, Farkas, & MacWhinney, 2004; Li & Zhao, 2013; Li, Zhao, & MacWhinney, 2007), the model highlights the dynamic interactions of L1 and L2 in the learning process, displays experience-dependent (input-dependent) plasticity and synaptic changes (in artificial neural networks), and accounts for age effects in both first and second languages by reference to principles of competition, entrenchment, self-organization, local and global reorganization, and Hebbian learning. These principles point to the effects of not only when L2 is learned, but also the manner in which the two languages interact, and the way early learning impacts later representation. Specifically, cross-language interconnections may be adapted differentially (e.g., as convergence or divergence), depending on the timing and extent of interaction and competition. Such adaptations occur as long as there is considerable flexibility and malleability in the networks of both language systems, in which the connection weights are not fully committed especially at early stages of learning. Importantly, the changes in the representation of the networks are directly due to learning itself, rather than to biological maturation from aging, as the model itself does not invoke different architectures at different times (though neurogenesis could be modeled as map growth in response to cognitive demands; see Li et al., 2004a, p. 1349). In short, the competition model examines how linguistic representations emerge dynamically out of interactions between the learning environment (features of the language to be learned) and the representation system (features of the learning brain), so that developmental trajectories can be clearly charted as a function of the interplay between experience and the underlying neural network (see Li, 2009, 2014; Li & Zhao, 2013; for further discussion).

Second, with regard to the *nature* of input, our review suggests that language experience is not unique in modifying structural properties of the brain, and cognitive skill acquisition in general (e.g., juggling, navigation, mirror reading) can similarly lead to structural neuroplasticity. This does not necessarily mean that linguistic and non-linguistic experiences are identical. There are indeed clear differences: one can learn the 3-ball juggling task within a few sessions, but one cannot learn a language in a day (despite commercial advertisements to the contrary). Our review shows that intensive juggling may modify structural properties of the brain in just 7 days (with even shorter timeframe reported for hippocampal volume changes due to spatial memory tasks),

but it takes about 3 months of intensive language training to see concomitant changes in anatomical structure due to language experience. This difference might be related to the nature, that is, the complexity of human language, which necessitates more efforts on the part of the learner for neuroanatomical changes to take place in the brain (see more discussion in Sections 3.3 and 3.4 on this difference).

Related to the effects of linguistic versus non-linguistic experience is the broader issue of whether experience gained in one domain can transfer to another domain or have a positive impact more generally. Fig. 3 and the results depicted seem to suggest that the training or the acquisition of a new skill is generally limited to anatomical changes in that specific brain region implicated for the relevant function (e.g., MT/V5 sensitive to motion perception in juggling). However, underlying the bilingual cognitive advantage hypothesis is the assumption that bilingual experience is somewhat unique: it is intensive, daily, long-term, and is practiced in high frequency for people living in a bilingual environment (Bialystok, Craik, & Luk, 2012; Kroll & Bialystok, 2013). As such, bilingual experiences may lead to cross-domain effects, resulting in an enhancement of both linguistic and non-linguistic, domain-general, functions rather than in a single-domain enhancement as in other cognitive skill acquisition situations (but see Bavelier et al., 2012 for how action video games may enhance resource allocation abilities and hence general cognition). The positive changes brought about by bilingual experience, especially on brain changes in structure and function as reviewed in this article, seem to point to broad cross-domain effects. This is particularly meaningful in light of the significant overlap in executive functions (e.g., switching, inhibiting, and monitoring) and language functions (e.g., phonological learning, lexical retrieval, morpho-syntactic processing), which rely on the same integrated brain network involving the IFG, ACC, IPL, and CN (Abutalebi & Green, 2007). Thus, it seems that both the complexity and the intensity of the bilingual experience could contribute to cross-domain transfer, as compared with non-linguistic experiences.

Third, with respect to the *extent* of the input, there is clear evidence from the bilingual studies that degree of anatomical changes (magnitude of GM density and WM integrity) may correlate with the performance level or proficiency level, and evidence from other cognitive domains is often, though not always, consistent with this correspondence, as reviewed. Such correlations could reflect the amount of experience of the learner, how often and to what degree the learner is exposed to the input, or the intensity that the input provides to the learner (which is related to the intensity of bilingual experience discussed above). Obviously, anatomical structure changes in a few brain regions alone will not allow the learner to use the second language as a native speaker, but if there is sufficient, consistent, and long-term, stimulation from the L2 across an extended period of the lifespan, brain changes will likely be accompanied by performance abilities that approximate or reach the skills of native speakers, and in exceptional cases, excelling above the average level of native-ness. Regardless of the actual final level of proficiency achieved, the extent of the specific linguistic experience in the L2 will be reflected in some combination of behavioral and neural response indices, which can be captured, as discussed above, increased or decreased

performance proficiency, neural activation, and anatomical changes (see discussion in Sections 3.3 and 3.4).

Several significant challenges to our current knowledge point to important future directions in the study of neuroplasticity induced by second language experience. One clear direction is that we need more systematic investigations into the complex function-structure-behavior relationships and their interactions (see discussion in Section 3.1). Although the structural imaging data as reviewed here are highly consistent with functional imaging patterns reported in the literature, there remain inconsistencies, in both linguistic and non-linguistic domains, as discussed. While most imaging findings, structural or functional, are also correlated with relevant performance variables, their correlations in some cases remain opaque, for example, sometimes positive and sometimes negative without clear reasons. This is especially worrisome when such correlations are considered against the underlying neurobiological basis of brain changes, in particular at the microscopic/microstructural level. As mentioned in Section 3.4, a number of researchers have called for efforts to identify the cellular and molecular mechanisms underlying experience-dependent structural changes related to learning and memory (see reviews in Kanai & Rees, 2011; May, 2011, and Zatorre et al., 2012), and this call applies equally to the study of language-induced brain changes. In other domains (e.g., spatial memory, motor behavior), one could conduct experiments *in vivo* using animals. In the case of language learning, this challenge becomes much more daunting given that it is difficult, if not impossible, to construct animal models of language.

A second area that has not received sufficient attention in the current literature is the impact of language typology (see Li, 2013, 2014). Language-specific characteristics vary widely in the degree to which they overlap or differ from one another; for example, grammatical morphology is a feature present in English and Spanish but absent in Chinese; Catalan, but not Spanish, has fine-grained front vowels as in English. It is reasonable to hypothesize that both short-term and long-term experiences with typologically dissimilar languages will yield different brain patterns, as compared to experiences with similar languages. In addition, larger cross-language overlap (more similarity) could lead to greater overlap in brain regions during the processing of the two languages, whereas smaller overlap (less similarity) could be associated with distinct neural response patterns in the bilingual's two languages (see Tolentino & Tokowicz, 2011 for review). For example, there have been recent functional neuroimaging data that indicate that typologically dissimilar languages such as Chinese and English may produce distinct patterns of neural responses in the bilingual's lexical representations (e.g., Chan et al., 2008; Yang, Tan, & Li, 2011). Although we know of no systematic comparison of anatomical differences as a function of language typology, there have been reports that experiences with language-specific characteristics (e.g., lexical tones in Chinese) may underlie some of the observed anatomical difference across cultures (e.g., Crinion et al., 2009; Kochunov et al., 2003). Future studies should explore the relationships among structure-function correspondences so as to reveal the impact of language typology more clearly.

A final promising direction is the understanding of the relationship between structural connectivity and functional

connectivity, particularly in light of recent trends in cognitive neuroscience to study brain networks of cognition (see Bressler & Menon, 2010; Friston, 2009; Menon, 2011; Sporns, 2011). This approach aims at integration (focus of new research direction) rather than segregation (focus of past neuroimaging research), because it examines not just individual brain regions, but also the spatial and temporal relationships between multiple brain regions during cognitive and linguistic tasks. In particular, Bressler and Menon (2010) provided a framework for analyzing brain networks for cognition, in which a set of sub-networks in the frontal, temporal, and parietal regions are identified to play key roles in attention, memory, and cognitive control. A number of studies have already used this framework to investigate functional brain networks of language learning (e.g., Sheppard et al., 2012; Yang & Li, 2012), but only two recent studies have examined structural brain networks for language in comparison with functional connectivity. As reviewed in Sections 2.2 and 2.3, Luk et al. (2011) compared WM structural connectivity with resting-state functional connectivity, and found parallel changes as a result of bilingual experience, particularly in the frontal regions. Garcia-Penton et al. (2014) conducted network-based analyses and identified two important sub-networks that had higher interconnectedness in bilinguals than in monolinguals. Currently, it remains unclear how the L1-L2 brain networks develop and change over time, and how such changes may manifest themselves differently in different individuals as a function of age, intensity of stimuli, and cognitive capacity before and after the learning experience. In particular, we need to understand the impact of L2 learning experience on the brain's frontal (e.g., IFG), temporal (e.g., TP), and striatal systems (e.g., CN), and how these subsystems interact with one another over time. Only large-scale, systematic, longitudinal, and long-term studies will allow us to address such questions in the future.

In conclusion, future research into brain networks, along with the study of individual characteristics of languages and individual differences of learners, will provide important pathways to a deeper understanding of the structure-function-behavior relationships in bilingualism and second language acquisition. It will also yield insights into the conditions under which L2 learning becomes successful and L2 experience-induced anatomical change becomes possible, and as such, the study of neuroplasticity as a function of second language learning has significant implications for our globalized society, as well as providing a window into the adaptive nature of the human mind and brain.

Acknowledgments

Cognition, Pennsylvania State University, University Park, PA 16802, USA. Email: pul8@psu.edu.

Appendix

Abbreviation	Brain Region
ACC	Anterior cingulate cortex
AC-OL	Anterior corpus callosum to the orbital lobe
AF	Arcuate fasciculus
AG	Angular gyrus
ATL	Anterior temporal lobe
ATP	Anterior temporal pole
ATR	Anterior thalamic radiation
BG	Basal ganglia
CB	Cerebellum
CC	Corpus callosum
CN	Caudate nucleus
DLPFC	Dorsolateral prefrontal cortex
DOL	Dorsal occipital lobe
EC	Extreme capsule
FX	Fornix
HG	Heschl's gyrus
HP	Hippocampus
IFG	Inferior frontal gyrus
IFGop	Inferior frontal gyrus pars opercularis
IFGpt	Inferior frontal gyrus pars triangularis
IFOF	Inferior fronto-occipital fasciculus
ILF	Inferior longitudinal fasciculus
IPL	Inferior parietal lobule
IPS	Inferior parietal sulcus
ITC	Inferior temporal cortex
ITG	Inferior temporal gyrus
LH	Left hemisphere
MFG	Middle frontal gyrus
mSFG	Medial superior frontal gyrus
MTG	Middle temporal gyrus
MT/V5	Middle temporal/visual area 5
POS	Parieto-occipital sulcus
RH	Right hemisphere
SFG	Superior frontal gyrus
SLF	Superior longitudinal fasciculus
SMA	Supplementary motor area
SMG	Supramarginal gyrus
SN	Substantia nigra
SOG	Superior occipital gyrus
SPG	Superior parietal gyrus
SPL	Superior parietal lobule
STG	Superior temporal gyrus
STP	Superior temporal pole
TP	Temporal pole
UF	Uncinate fasciculus
V2/3	Visual area 2 and 3

REFERENCES

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica, 128*(3), 466—478.

Abutalebi, J., Canini, M., Della Rosa, P. A., Sheung, L. P., Green, D. W., & Weekes, B. S. (2014). Bilingualism protects anterior temporal lobe integrity in aging. *Neurobiology of Aging, 35*(9), 2126—2133.

Abutalebi, J., Cappa, S. F., & Perani, D. (2005). What can functional neuroimaging tell us about the bilingual brain? In J. F. Kroll, & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 497–515) New York: Oxford University Press.

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., & Keim, R. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex, 22*(9), 2076–2086.

Abutalebi, J., & Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *Journal of Neurolinguistics, 20*(3), 242–275.

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics, 4*(3), 316–329.

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., & Shukla, A. K. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurocase, 81*, 1–7.

Antoniou, M., Gunasekera, G. M., & Wong, P. C. M. (2013). Foreign language training as cognitive therapy for age-related cognitive decline: a hypothesis for future research. *Neuroscience & Biobehavioral Reviews, 37*(10), 2689–2698.

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—The methods. *NeuroImage, 11*(6), 805–821.

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.

Aydin, K., Ucar, A., Oguz, K. K., Okur, O. O., Agayev, A., Unal, Z., et al. (2007). Increased gray matter density in the parietal cortex of mathematicians: a voxel-based morphometry study. *American Journal of Neuroradiology, 28*(10), 1859–1864.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Bates, E. (1999). Plasticity, localization and language development. In S. H. Broman, & J. Fletcher (Eds.), *The changing nervous system: Neurobehavioral consequences of early brain disorders*. New York: Oxford University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language aquisition*. Hillsdale, NJ: Lawrence Erlbaum.

Bavelier, D., Green, C. S., Pouget, A., & Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annual Review of Neuroscience, 35*(1), 391–416.

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*, 3–12.

Bialystok, E., & Barac, R. (2013). Cognitive effects. In F. Grosjean, & P. Li (Eds.), *The psycholinguistics of bilingualism*. New York, NY: John Wiley & Sons, Inc.

Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459–464.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences, 16*(4), 240–250.

Binder, J., & Desai, R. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*, 527–536.

Boyke, J., Driemeyer, J., Gaser, C., Büchel, C., & May, A. (2008). Training-induced brain structure changes in the elderly. *Journal of Neuroscience, 28*(28), 7031–7035.

Bressler, S. L., & Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends in Cognitive Sciences, 14*(6), 277–290.

Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., et al. (2009). An anatomical signature for literacy. *Nature, 461*, 983–986.

Chan, A., Luke, K., Li, P., Yip, V., Li, G., Weekes, B., et al. (2008). Neural correlates of nouns and verbs in early bilinguals. *Annals of the New York Academy of Sciences, 1145*, 30–40.

Chung, M. K., Dalton, K. M., Shen, L., Evans, A. C., & Davidson, R. J. (2006). *Unified cortical surface morphometry and its application to quantifying amount of gray matter*. Technical Report No. 1122. Department of Statistics, U. of Wisconsin-Madison.

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

Crinion, J. T., Green, D. W., Chung, R., Ali, N., Grogan, A., Price, G. R., et al. (2009). Neuroanatomical markers of speaking Chinese. *Human Brain Mapping, 30*, 4108–4115.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., et al. (2006). Language control in the bilingual brain. *Science, 312*(5779), 1537–1540.

Cummine, J., & Boliek, C. A. (2013). Understanding white matter integrity stability for bilinguals on language status and reading performance. *Brain Structure & Function, 218*(2), 595–601.

Della Rosa, P. A., Videsott, G., Borsa, V. M., Canini, M., Weekes, B. S., Franceschini, R., et al. (2013). A neural interactive location for multilingual talent. *Cortex, 49*(2), 605–608.

Diamond, J. (2010). The benefits of multilingualism. *Science, 330*, 332–333.

Dong, Y., & Li, P. (2014). The cognitive science of bilingualism. In T. Swaab (Ed.), *Language and Linguistics Compass* (in press).

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, N.J: L. Erlbaum.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bodgahn, U., & May, A. (2004). Changes in grey matter induced by training. *Nature, 427*, 311–312.

Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., & May, A. (2008). Changes in gray matter induced by learning—Revisited. *PLoS One, 3*(7), e2669.

Elman, J. (1993). Learning and development in neural networks: the importance of starting small. *Cognition, 48*, 71–99.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Elmer, S., Hänggi, J., & Jäncke, L. (2014). Processing demands upon cognitive, linguistic, and articulatory functions promote grey matter plasticity in the adult multilingual brain: Insights from simultaneous interpreters. *Cortex*, 1–33.

Elmer, S., Hänggi, J., Meyer, M., & Jäncke, L. (2011). Differential language expertise related to white matter architecture in regions subserving sensory-motor coupling, articulation, and interhemispheric transfer. *Human Brain Mapping, 32*(12), 2064–2074.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., et al. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 3017–3022.

Filler, A. (2009). MR neurography and diffusion tensor imaging: origins, history & clinical impact of the first 50,000 cases with an assessment of efficacy and utility in a prospective 5,000 patient study group. *Neurosurgery*, A29–A43.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences, 97*(20), 11050–11055.

Flöel, A., de Vries, M. H., Scholz, J., Breitenstein, C., & Johansen-Berg, H. (2009). White matter integrity in the vicinity of Broca's area predicts grammar learning success. *NeuroImage, 47*(4), 1974–1981.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. The MIT Press.

Friston, K. J. (2009). Modalities, modes, and models in functional neuroimaging. *Science, 326*, 399–403.

Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academy of Sciences, 95*(3), 906—913.

Garcia-Penton, L., Perez, F. A., Iturria-Medina, Y., Gillon-Dowens, M., & Carreiras, M. (2014). Anatomical connectivity changes in the bilingual brain. *NeuroImage, 84*, 495—504.

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.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex, 17*(3), 575—582.

Golestani, N., & Pallier, C. (2007). Anatomical correlates of foreign speech sound production. *Cerebral Cortex, 17*(4), 929—934.

Golestani, N., Paus, T., & Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron, 35*, 997—1010.

Grogan, A., Green, D. W., Ali, N., Crinion, J. T., & Price, C. J. (2009). Structural correlates of semantic and phonemic fluency ability in first and second languages. *Cerebral Cortex, 19*(11), 2690—2698.

Grogan, A., Jones, Ō. P., Ali, N., Crinion, J., Orabona, S., Mechias, M. L., et al. (2012). Structural correlates for lexical efficiency and number of languages in non-native speakers of English. *Neuropsychologia, 50*(7), 1347—1352.

Grosjean, F., & Li, P. (2013). *The psycholinguistics of bilingualism.* New York, NY: John Wiley & Sons, Inc.

Hernandez, A. E. (2013). *The bilingual brain.* Oxford, UK: Oxford University Press.

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638—650.

Hernandez, A. E., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences, 9*, 220—225.

Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews, 6*(3), 121—143.

Hickok, G., & Poeppel, D. (2007). Cortical organization of speech processing. *Nature Review Neuroscience, 8*, 393—402.

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., & Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: a multimodal imaging study. *Journal of Neuroscience, 33*(34), 13663—13672.

Ilg, R., Wohlschlager, A. M., Gaser, C., Liebau, Y., Dauner, R., & Woller, A. (2008). Gray matter increase induced by practice correlates with task-specific activation: a combined functional and morphometric magnetic resonance imaging study. *Journal of Neuroscience, 28*(16), 4210—4215.

Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: which suggested differences can we trust and what do they mean? *Language Learning, 56*(1), 279—304.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60—99.

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience, 12*(4), 231—242.

Kennedy, D., & Norman, C. (2005). What don't we know. *Science, 309*(5731), 78—102.

Kim, K. H. S., Relkin, N. R., Lee, K. M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature, 388*, 171—175.

Kim, J. S., Singh, V., Lee, J. K., Lerch, J., Ad-Dab'bagh, Y., & MacDonald, D. (2005). Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *NeuroImage, 27*(1), 210—221.

Klein, D., Mok, K., Chen, J.-K., & Watkins, K. E. (2013). Age of language learning shapes brain structure: a cortical thickness study of bilingual and monolingual individuals. *Brain and Language*, 1—5.

Kochunov, P., Fox, P., Lancaster, J., Tan, L. H., Amunts, K., Zilles, K., et al. (2003). Localized morphological brain differences between English-speaking Caucasians and Chinese-speaking Asians: new evidence of anatomical plasticity. *NeuroReport, 14*, 961—964.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology, 25*(5), 497—514.

Kunimatsu, A., Aoki, S., Masutani, Y., Abe, O., Hayashi, N., Mori, H., et al. (2004). The optimal trackability threshold of fractional anisotropy for diffusion tensor tractography of the corticospinal tract. *Magnetic Resonance in Medical Sciences, 3*, 11—17.

Kwok, V., Niu, Z., Kay, P., Zhou, K., Mo, L., Jin, Z., et al. (2011). Learning new color names produces rapid increase in gray matter in the intact adult human cortex. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 6686—6688.

Lee, H., Devlin, J. T., Shakeshaft, C., Stewart, L. H., Brennan, A., & Glensman, J. (2007). Anatomical traces of vocabulary acquisition in the adolescent Brain. *Journal of Neuroscience, 27*(5), 1184—1189.

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

Lerch, J. P., & Evans, A. C. (2005). Cortical thickness analysis examined through power analysis and a population simulation. *NeuroImage, 24*(1), 163—173.

Li, P. (2009). Lexical organization and competition in first and second languages: Computational and neural mechanisms. *Cognitive Science, 33*, 629—664.

Li, P. (2013). The cross-cultural bilingual brain: comment on "The bilingual brain: flexibility and control in the human cortex" by Buchweitz and Prat. *Physics of Life Reviews, 10*, 446—447.

Li, P. (2014). Bilingualism as a dynamic process. In B. MacWhinney, & W. O'Grady (Eds.), *Handbook of language emergence*. John Wiley & Sons, Inc (in press).

Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks, 17*, 1345—1362.

Li, P., & Green, D. (2007). Neurocognitive approaches to bilingualism: Asian languages. *Bilingualism: Language and Cognition, 10*, 117—119.

Li, P., Jin, Z., & Tan, L. H. (2004). Neural representations of nouns and verbs in Chinese: an fMRI study. *NeuroImage, 21*(4), 1533—1541.

Li, P., & MacWhinney, B. (2013). Competition model. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Malden, MA: John Wiley & Sons, Inc.

Li, P., & Tokowicz, N. (2012). The psycholinguistics of second language acquisition. In P. Robinson (Ed.), *Routledge encyclopedia of second language acquisition*. New York, NY: Taylor & Francis/Routledge.

Li, P., & Zhao, X. (2013). Self-organizing map models of language acquisition. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00828. Article #828.

Li, P., Zhao, X., & MacWhinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive Science, 31*, 581—612.

Li, Y., Yang, J., Scherf, S., & Li, P. (2013). Two faces, two languages: an fMRI study of picture naming in bilinguals. *Brain and Language, 3*, 452—462.

Loui, P., Li, H. C., & Schlaug, G. (2011). White matter integrity in right hemisphere predicts pitch-related grammar learning. *NeuroImage, 55*(2), 500—507.

Lövdén, M., Wenger, E., Mårtensson, J., Lindenberger, U., & Bäckman, L. (2013). Structural brain plasticity in adult learning and development. *Neuroscience and Biobehavioral Reviews, 37*(9), 2296—2310.

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.

MacWhinney, B. (2012). The logic of the Unified Model. In S. Gass, & A. Mackey (Eds.), *Handbook of second language acquisition*. New York: Routledge.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences, 97*(8), 4398—4403.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S. J., & Burgess, N. (2003). Navigation expertise and the human hippocampus: a structural brain imaging analysis. *Hippocampus, 13*(2), 250—259.

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus, 16*(12), 1091—1101.

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., et al. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage, 63*(1), 240—244.

Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science, 270*(5233), 102—105.

May, A. (2011). Experience-dependent structural plasticity in the adult human brain. *Trends in Cognitive Sciences, 15*(10), 475—482.

May, A., & Gaser, C. (2006). Magnetic resonance-based morphometry: a window into structural plasticity of the brain. *Current Opinion in Neurology, 19*(4), 407—411.

May, A., Hajak, G., Ganssbauer, S., Steffens, T., Langguth, B., Kleinjung, T., et al. (2007). Structural brain alterations following 5 days of intervention: dynamic aspects of neuroplasticity. *Cerebral Cortex, 17*(1), 205—210.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J. P., Ashburner, J., Frackowiak, R. S., et al. (2004). Structural plasticity in the bilingual brain: proficiency in a second language and age at acquisition affect grey-matter density. *Nature, 431*(7010), 757.

Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel-based morphometry of the human brain: methods and applications. *Current Medical Imaging Reviews, 1*(1), 1—9.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences, 15*(10), 483—506.

Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: working memory as language aptitude. In A. Healy, & L. Bourne (Eds.), *Foreign language learning: Psycholinguistics studies on training and retention*. Mahwah, NJ: Lawrence Erlbaum.

Mohades, S. G., Struys, E., Van Schuerbeek, P., Mondt, K., Van Den Craen, P., & Luypaert, R. (2012). DTI reveals structural differences in white matter tracts between bilingual and monolingual children. *Brain Research, 1435*, 72—80.

Parker Jones, Ō., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., & Price, C. J. (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cerebral Cortex, 22*(4), 892—902.

Pliatsikas, C., Johnstone, T., & Marinis, T. (2013). Grey matter volume in the cerebellum is related to the processing of grammatical rules in a second language: a structural voxel-based morphometry study. *Cerebellum*, 1—9.

Poeppel, D., Emmorey, K., Hickok, G., & Pylkkänen, L. (2012). Towards a new neurobiology of language. *The Journal of Neuroscience, 32*(41), 14125—14131.

Price, C. J. (2000). The anatomy of language: contributions from functional neuroimaging. *Journal of Anatomy, 197*, 335—359.

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences, 1191*(1), 62—88.

Ressel, V., Pallier, C., Ventura-Campos, N., Diaz, B., Roessler, A., Ávila, C., et al. (2012). An effect of bilingualism on the auditory cortex. *Journal of Neuroscience, 32*(47), 16597—16601.

Richardson, F. M., & Price, C. J. (2009). Structural MRI studies of language function in the undamaged brain. *Brain Structure & Function, 213*(6), 511—523.

Roberts, R. E., Bain, P. G., Day, B. L., & Husain, M. (2013). Individual differences in expert motor coordination associated with white matter microstructure in the cerebellum. *Cerebral Cortex, 23*, 2282—2292.

Rodriguez-Fornells, A., Cunillera, T., Mestres-Misse, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1536), 3711—3735.

Sagi, Y., Tavor, I., Hofstetter, S., Tzur-Moryosef, S., Blumenfeld-Katzir, T., & Assaf, Y. (2012). Learning in the fast lane: new insights into neuroplasticity. *Neuron, 73*(6), 1195—1203.

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.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience, 5*, 688—694.

Scholz, J., Klein, M. C., Behrens, T. E. J., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nature Neuroscience, 12*(11), 1370—1371.

Sebastián-Gallés, N., Soriano-Mas, C., Baus, C., Díaz, B., Ressel, V., & Pallier, C. (2012). Neuroanatomical markers of individual differences in native and non-native vowel perception. *Journal of Neurolinguistics, 25*(3), 150—162.

Sheppard, J. P., Wang, J. P., & Wong, P. C. (2012). Large-scale cortical network properties predict future sound-to-word learning success. *Journal of Cognitive Neuroscience, 24*(5), 1087—1103.

Spalding, K. L., Bergmann, O., Alkass, K., Bernard, S., Salehpour, M., Huttner, H. B., et al. (2013). Dynamics of hippocampal neurogenesis in adult humans. *Cell, 153*(6), 1219—1227.

Sporns, O. (2011). *Networks of the brain*. Cambridge, MA: MIT Press.

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., et al. (2012). Structural plasticity in the language system related to increased second language proficiency. *Cortex, 48*(4), 458—465.

Sutherland, M. E., Zatorre, R. J., Watkins, K. E., Hervé, P.-Y., Leonard, G., & Pike, B. G. (2012). Anatomical correlates of dynamic auditory processing: relationship to literacy during early adolescence. *NeuroImage, 60*(2), 1287—1295.

Thomas, C., & Baker, C. I. (2013). Teaching an adult brain new tricks: a critical review of evidence for training-dependent structural plasticity in humans. *NeuroImage, 73*, 225—236.

Tolentino, L. C., & Tokowicz, N. (2011). Across languages, space, and time: a review of the role of cross-language similarity in L2 (morpho)syntactic processing as revealed by fMRI and ERP. *Studies in Second Language Acquisition, 33*, 1—34.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*(1), 38—69.

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition, 92*(1—2), 231—270.

van Hell, J. G., & Tokowicz, N. (2010). Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research, 26*(1), 43—74.

Ventura-Campos, N., Sanjuan, A., Gonzalez, J., Palomar-Garcia, M. A., Rodriguez-Pujadas, A., & Sebastián-Gallés, N. (2013). Spontaneous brain activity predicts learning ability of foreign sounds. *Journal of Neuroscience, 33*(22), 9295—9305.

Veroude, K., Norris, D. G., Shumskaya, E., Gullberg, M., & Indefrey, P. (2010). Functional connectivity between brain regions involved in learning words of a new language. *Brain and Language, 113*(1), 21—27.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron, 37*, 159—170.

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience, 8*(3), 231—256.

Wong, P. C. M., Perrachione, T. K., & Parrish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping., 1006*, 995—1006.

Wong, P. C. M., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., et al. (2008). Volume of left heschl's gyrus and linguistic pitch learning. *Cerebral Cortex, 18*(4), 828—836.

Xiang, H., Dediu, D., Roberts, L., van Oort, E., Norris, D. G., & Hagoort, P. (2012). The structural connectivity underpinning language aptitude, working memory, and IQ in the perisylvian language network. *Language Learning, 62*(Suppl. 2), 110—130.

Yang, J., Gates, K., Molenaar, P., & Li, P. (2014). *Neural changes underlying successful second language word learning: An fMRI study*. Manuscript submitted for publication.

Yang, J., & Li, P. (2012). Brain networks of explicit and implicit learning. *PLoS One, 7*(8), e42993.

Yang, J., Tan, L., & Li, P. (2011). Lexical representation of nouns and verbs in the late bilingual brain. *Journal of Neurolinguistics, 24*, 674—682.

Zatorre, R. J., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in gray and white: neuroimaging changes in brain structure during learning. *Nature Neuroscience, 15*(4), 528—536.

Zhang, L., Xi, J., Xu, G., Shu, H., Wang, X., & Li, P. (2011). Cortical dynamics of acoustic and phonological processing in speech perception. *Plos One, 6*, 6.

Zou, L., Ding, G., Abutalebi, J., Shu, H., & Peng, D. (2012). Structural plasticity of the left caudate in bimodal bilinguals. *Cortex, 48*(9), 1197—1206.