The effect of lifelong bilingualism on regional grey and white matter volume

Running head: BILINGUALISM AND BRAIN STRUCTURE

The effect of lifelong bilingualism on regional grey and white matter volume

Rosanna K. Olsen*[1]
Melissa M. Pangelinan*[1]
Cari Bogulski[2]
M. Mallar Chakravarty[1, 3, 4, 5]
Gigi Luk[6]
Cheryl Grady[1, 7, 8]
Ellen Bialystok [1, 2]

Affiliation:
1 Rotman Research Institute, Baycrest Health Sciences, 3560 Bathurst Street, Toronto, ON M6A 2E1, Canada
2 Department of Psychology, York University, 4700 Keele St. Toronto, ON M3J 1P3, Canada
3 Cerebral Imaging Centre, Douglas Mental Health University Institute, 6875 LaSalle Boulevard, Montreal, QC H4H 1R3, Canada
4 Department of Psychiatry, McGill University, 1033 Pine Avenue West, Montreal, QC H3A 1A1, Canada
5 Department of Biomedical Engineering, McGill University, 3775 Rue University, Montréal, QC H3A 2B4, Canada
6 Harvard Graduate School of Education, Appian Way, Cambridge, MA 02138, USA
7 Department of Psychology, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada
8 Department of Psychiatry, University of Toronto, 250 College Street, Toronto, ON M5T 1R8, Canada

*These authors contributed equally to this work
Corresponding author: Ellen Bialystok, York University, 4700 Keele St. Toronto, ON M3J 1P3, Canada, ellenb@yorku.ca

Abstract

Lifelong bilingualism is associated with the delayed diagnosis of dementia, suggesting bilingual experience is relevant to brain health in aging. While the effects of bilingualism on cognitive functions across the lifespan are well documented, less is known about the neural substrates underlying differential behavior. It is clear that bilingualism affects brain regions that mediate language abilities and that these regions are at least partially overlapping with those that exhibit age-related decline. Moreover, the behavioural advantages observed in bilingualism are generally found in executive function performance, suggesting that the frontal lobes may also be sensitive to bilingualism, which exhibit volume reductions with age. The current study investigated structural differences in the brain of lifelong bilingual older adults (n = 14, mean age = 70.4) compared with older monolinguals (n = 14, mean age= 70.6). We employed two analytic approaches: 1) we examined global differences in grey and white matter volumes; and, 2) we examined local differences in volume and cortical thickness of specific regions of interest previously implicated in bilingual/monolingual comparisons (temporal pole) or in aging (entorhinal cortex and hippocampus). We expected bilinguals would exhibit greater volume of the frontal lobe and temporal lobe (grey and white matter), given the importance of these regions in executive and language functions, respectively. We further hypothesized that regions in the medial temporal lobe, which demonstrate early changes in aging and exhibit neural pathology in dementia, would be more preserved in the bilingual group. As predicted, bilinguals exhibit greater frontal lobe white matter compared with monolinguals. Moreover, increasing age was related to decreasing temporal pole cortical thickness in the monolingual group, but no such relationship was observed for bilinguals. Finally, Stroop task performance was positively correlated with frontal lobe white matter, emphasizing the importance of preserved white matter in maintaining executive function in aging. These results underscore previous findings implicating an association between bilingualism and preserved frontal and temporal lobe function in aging.

Key words: bilingual, cognitive reserve, MRI, volumetric, aging, hippocampus

I. Introduction

Aging is associated with significant reductions in both global and regional brain volumes. In grey matter, the trajectory of these age-related changes is variable across the brain such that more rapid decline is observed in regions of the frontal and parietal lobes than in the temporal and occipital lobes (Fjell et al., 2013; Kennedy & Raz, 2009; Raz et al., 2004, 2005; Resnick et al., 2000; Resnick et al., 2003; Walhovd et al., 2005, 2011). In white matter, some studies have reported a more diffuse pattern of volumetric reductions across the brain with age (Fjell et al., 2013; Resnick et al., 2000, 2003; Salat et al., 2009; Walhovd et al., 2005, 2011), whereas others have reported greater reduction in frontal lobe white matter compared to other regions (Brickman et al., 2006; Ferreira et al., 2014; Head et al., 2004). In addition to neocortical changes, it is well documented that the hippocampus decreases in volume with increasing age (Jack et al., 1998; Mueller et al., 1998; Resnick et al., 2003; Walhovd et al., 2005). These structural differences in the frontal lobes and hippocampus are associated with corresponding changes in behavioural functions, with significant age-related differences observed for executive and memory functions (Brickman et al., 2006; Cardenas et al., 2011; Kennedy & Raz, 2009; Petersen et al., 2000).

Longitudinal analyses of age-related decline in brain volume have revealed large heterogeneity in the magnitude of tissue reductions amongst healthy older adults (Harada, Natelson Love, & Triebel, 2013; Resnick et al., 2003), suggesting that environmental factors and/or genetics may alter the trajectory of change in the brains of older adults. Indeed, lifestyle factors such as engagement in exercise (Erickson et al., 2011; Kramer & Erickson, 2007), playing a musical instrument (Hanna-Pladdy & MacKay, 2011; Schlaug et al., 1995; Wan & Schlaug, 2010), and speaking a second language (Bialystok et al., 2004; Bialystok, Craik, & Ryan, 2006; Salvatierra & Rosselli, 2010) appear to stave off age-related declines in cognitive functions, especially those requiring executive control (e.g., working memory, inhibition, task-switching, etc.). In addition, corresponding increases in brain volume due to engagement in aerobic exercise (Ahlskog, Geda, Graff-Radford, & Petersen, 2011; Colcombe et al., 2006; Erickson et al.,

2011) and playing a musical instrument (Gaser & Schlaug, 2003; Zatorre, Fields, & Johansen-Berg, 2012) have been reported. There is an emerging literature that has examined structural differences in the brain with respect to learning a second language (e.g., Mårtensson et al., 2012; Mechelli et al., 2004; Schlegel, Rudelson, & Tse, 2012; Stein et al., 2012). For example, following immersion in a second language, young adults demonstrated increases in grey matter in the inferior frontal gyrus and anterior temporal lobe (Stein et al., 2012). Similarly, immersive language training resulted in greater thickness in the regions of the frontal lobe and superior temporal gyrus, and larger volume in the hippocampus (Mårtensson et al., 2012). However, much less is known about how lifelong bilingualism (as opposed to short-term language training) affects grey and white matter brain structure (Gold et al., 2013a; see Stein et al., 2014, for a review).

The behavioural consequences associated with lifelong bilingualism have been extensively reported (Bialystok et al., 2009; Bialystok, Craik, & Luk, 2012). Specifically, bilinguals outperform monolinguals on tasks that rely on various aspects of frontally-mediated executive functions (Bialystok et al., 2004; Bialystok et al., 2006; Bialystok et al., 2014; Salvatierra and Rosselli 2010; Schroeder and Marian 2012; Gold et al. 2013b) and language control mediated by the coordination of frontal and temporal lobes (Luk et al., 2012); this bilingual advantage is especially pronounced in older participants. Such cognitive benefits are thought to have profound and long-lasting consequences on brain health, as evidenced by a delay in the onset of Alzheimer's-type dementia in older bilingual adults relative to monolinguals of the same age (Alladi et al., 2013; Bialystok et al., 2014; Bialystok, Craik, & Freedman, 2007; Wilson, et al., in press). Yet, little is known about the effect of lifelong bilingualism, as a natural life experience, on brain structure, particularly in older adults (Luk, Bialystok, Craik, & Grady, 2011; see Bialystok, Craik, & Luk, 2012 for a review). Recently, Abutalebi and colleagues (2014) found greater grey matter volumes in the left anterior temporal pole, a region known for its role in semantic processing (Bonner & Price, 2013) and the retrieval of proper names (Ross et al., 2010) in older bilinguals relative to monolinguals. This same study also demonstrated significant group differences in the relationship between temporal

pole grey matter and age, suggesting that bilingualism may protect this region from age-related structural decline.

In a previous study we used diffusion tensor imaging to investigate differences in white matter microstructure in the same sample as the present study. We found that compared to older monolinguals, older bilinguals exhibited greater fractional anisotropy, a measure of anisotropic water diffusion, in the corpus callosum, as well as the superior and inferior longitudinal fascicule (Luk et al. 2011). Volumetric measurements of white matter (reported here) and DTI (reported previously) can be thought of as complementary neuroimaging techniques, both approaches have strengths and weaknesses (Jones et al., 2013). Volumetric measurements of grey and white matter volume and cortical thickness have been examined extensively across the lifespan (Walhovd et al., 2005; Fjell et al., 2009; Giedd and Rapoport, 2010), have been demonstrated to be powerful neurodevelopmental phenotypes (Lenroot et al., 2007, 2009), and are influenced strongly by age and environment and lifestyle factors (Nyberg et al., 2012).

T1-weighted structural MRI is optimized for the segmentation of different tissue types (grey matter, white matter, and cerebral spinal fluid), and as such is ideal for volumetric assessments of both grey matter and white matter brain structures. Therefore, in the current study, we used T1-weighted MRI to estimate volumetric differences between lifelong older bilinguals (n = 14, *M* age = 70.6 years, *SD* = 3 years) and monolinguals of the same age (n = 14; *M* age = 70.4 years, *SD* = 3.7 years). In addition to group-level differences (i.e., group main effects), we also examined if lifelong bilinguals exhibited a different relationship between structure and age than monolinguals (i.e., Age x Group interactions). The latter analysis was used to investigate whether certain brain regions demonstrated a pattern consistent with differential age-related atrophy in the two groups. We quantitatively assessed grey and white matter brain structure in older adults using two analytic approaches: 1) we examined global differences in grey and white matter lobar volumes (frontal, parietal, temporal, and occipital); and, 2) we examined

local differences in specific regions of interest (ROIs) within the temporal lobe. We focused on the structures within the temporal lobe that are especially affected by aging and dementia (entorhinal cortex and hippocampus) as well as a region which is involved in language processing and was previously implicated in bilingual/monolingual comparisons within a similarly aged cohort (temporal pole; Abutalebi et al., 2014). For the ROI analysis, we examined the volume of the entorhinal cortex, hippocampus, and temporal pole. We also examined cortical thickness for the entorhinal cortex and the temporal pole; it is not common practice to compute thickness for the hippocampus using automated procedures on standard resolution (1x1x1mm voxel size) structural MRI data. The combined use of these two approaches allowed for the examination of more distributed or global differences in brain structure at the level of lobar differences in grey and white matter as well as a targeted examination of temporal lobe ROIs in which we expected to observe group differences. The temporal pole is a region of particular interest given that is affected by normal aging as well as semantic dementia, and is thought to play a critical role in lexical retrieval (Tranel, 2009). Given the benefit of bilingualism on behavioural functions, we predicted that compared to monolinguals, lifelong bilinguals would exhibit greater volume and/or reduced age-related differences in brain regions that support executive function and language processing, such as the frontal and temporal lobes. We also predicted that lifelong bilingualism would be associated with greater volume and/or reduced age-related differences in the medial temporal lobe, a region that has structural sensitivity to aging.

2. Results

2.1 Volumetric analysis of neocortical structures

2.1.1 Total intracranial volume

Intracranial volume (ICV) was calculated for each participant and was used to correct for differences in head size (i.e., left frontal white matter $volume_{norm}$= left frontal white matter / ICV). As such, it was first critical to establish that there were no group differences in ICV in order to determine that using this

measurement to normalize the neocortical and hippocampal volumes, as presented below, did not introduce group biases absent from the raw data. As expected, ICV was equivalent across groups (t = 0.05, p > 0.05).

2.1.2 Grey matter volume

To investigate the effect of lifelong bilingualism on grey and white matter brain volume, repeated measures ANOVA was conducted with lobe (frontal, temporal, parietal, occipital) and hemisphere (left, right) entered as within-subject factors and language group (monolingual, bilingual) entered as a between-subject factor. The first analysis focused on the effect of language group on grey matter volume within the four major lobes of the brain. This analysis revealed no significant effect of language group on grey matter $volume_{norm}$ overall (F = 0.24, p > 0.05) and this between-subjects factor did not interact with lobe or hemisphere (ps > 0.05 for all).

We next examined whether different relationships between overall lobar grey matter volume and age existed across the two groups. The age-based regression coefficients for each grey matter lobar region were compared between monolinguals and bilingual groups. Because group by hemisphere interactions were not observed in the main analysis above, we collapsed left and right hemispheres. No group differences in the age-related slopes were observed for any grey matter lobar region for frontal, parietal, temporal, and occipital (Fs < 1.1, ps > 0.05 for all).

2.1.3 White matter volume

The second analysis examined the effects of language experience on white matter volume within the four major lobes of the brain. As described above, repeated measures ANOVA was conducted with lobe and hemisphere entered as within-subject factors and language group (monolingual, bilingual) was entered as a between-subject factor. The results from this analysis revealed no significant effect of language experience on white matter $volume_{norm}$ across the entire brain (no main effect when analyzed

collapsed across lobes, $F = 1.47$, $p > 0.05$); however, a significant language group by lobe interaction was observed ($F = 2.84$, $p = 0.04$).

The significant interaction between language group and lobe indicated that the magnitude of the difference between monolinguals and bilinguals in white matter $volume_{norm}$ depended on lobar region (Figure 1; Table 2). To investigate this interaction, follow-up comparisons were conducted by assessing volumetric differences across the two groups for each lobe separately. Because hemisphere did not interact significantly with group, white matter volumes from the two hemispheres were added into a single measure and then normalized by ICV. These comparisons revealed that white matter $volume_{norm}$ differed significantly between language groups (bilinguals > monolinguals) in the frontal lobe ($t = 1.67$, $p = 0.05$), marginally differed in the temporal lobe, $p = 0.09$); but did not differ within the parietal ($p > 0.10$) or occipital lobes ($p > 0.10$).

To investigate if different age-related trajectories were found across the two groups, we examined the regression coefficients for each white matter lobar region. No significant group differences in the age-related slopes were observed among the white matter lobar regions ($F = 1.73, 0.92, 2.86, 0.01$ for frontal, parietal, temporal, and occipital, $p > 0.05$ for all).

2.2 Regional analysis of the hippocampus, entorhinal cortex, and temporal pole

No mean group differences were observed for the volume of the hippocampus (Figure 2), entorhinal cortex, or temporal pole ($F = 0.67, 0.04, 0.56$, respectively, $p > 0.05$ for all). Similarly, no mean group differences were observed for cortical thickness of the entorhinal cortex or temporal pole ($F = 0.06$ and 0.84, respectively, ps > 0.05).

We next investigated if the two groups differed with respect to age-related trajectories in these regions of interest (i.e., Group x Age interaction). No differences in the age-related slopes were observed for

the volume of the hippocampus, entorhinal cortex and temporal pole (F = 0.06, 1.26, and 1.54, respectively, $p > 0.05$ for all) or for the entorhinal thickness (F = 2.78, $p > 0.05$). However, a group by age interaction was observed for temporal pole thickness (F = 6.6, p = 0.02; Figure 3). In addition, a significant negative relationship between age and temporal pole thickness was observed for monolinguals (β = -0.018, SE = 0.007, t = -2.61, p = 0.02, R^2 = 0.34), but not for bilinguals (β = 0.007, SE = 0.007, t = 1.03, $p > 0.05$, R^2 = 0.10).

2.3 Brain-behaviour relationships

The relationship between brain volume and Stroop performance (which was equivalent across groups, see Table 1) was examined to provide supporting evidence for the involvement of frontal lobe white matter in resolving interference. We first observed that across the two groups, better Stroop performance (i.e. less interference or a smaller difference in response time between the interference condition and baseline performance) was associated with larger white matter volume in the frontal lobes (Figure 4A; r = - 0.39, p = 0.03). When this relationship was probed in each group separately, a significant relationship was observed in the bilingual group (Figure 4B, r = - 0.52, p = 0.03) but was not statistically significant in the monolingual group (r = - 0.32, p = 0.15). Fisher r-to-z transformation revealed that the correlation coefficients do not differ significantly across groups (z = - 0.55, p = 0.58), which suggests that the relationship between Stroop performance and white matter volume does not differ appreciably between the two groups.

3. Discussion

3.1 Summary

The current study examined whether individuals with lifelong bilingual experience exhibit greater neocortical grey and white matter lobar volumes and grey matter structure in targeted temporal lobe regions than monolingual individuals. Bilinguals exhibited significantly greater frontal lobe and marginally greater temporal lobe white matter volumes than monolinguals of the same age and

comparable demographic backgrounds. These results suggest that lifelong bilingualism may result in enhanced neuronal connections amongst brain areas important for higher-order cognitive and language functions. These regions of enhanced white matter volume in bilinguals correspond to the cognitive functions for which bilingual advantages have been reported (e.g., executive function). Indeed, Stroop task interference performance, a classic assessment of response inhibition, was significantly correlated with frontal lobe white matter (a relationship that has also been observed in young adults, see Takeuchi et al., 2012). Furthermore, we observed a group difference in the relationship between cortical thickness and age within the temporal lobe, specifically, in the temporal pole. These data, together with previous findings in young and older adults, suggests that lifelong bilingualism may increase regional brain volume throughout the lifespan and/or attenuate age-related reductions in older adults in brain regions critical for language (Figure 5).

3.2 Structural brain changes associated with lifelong bilingualism

We found significant group differences in frontal white matter volume as well as a different relationship between age and temporal pole grey matter thickness in bilingual older adults than in monolinguals. With respect to the grey matter differences, our results are consistent with previous studies in young adults in that second language learning may affect the trajectory of grey matter structure (thickness or volume) in regions of the temporal lobe (Mårtensson et al., 2012; Stein et al., 2012). Such differences in grey matter volume of the temporal pole as a function of bilingualism have previously been reported for older adults (Abutalebi et al., 2014). Taken together, it appears that lifelong bilingualism may result in a different trajectory for grey matter development in a region known for its role in semantic processing, semantic retrieval, and naming.

With respect to the white matter, these results are consistent with a recent study employing diffusion tensor imaging to assess white matter microstructure (e.g., FA); FA was greater in young adults with second-language training compared to controls in tracts within the frontal and temporal lobes (Schlegel

et al., 2012). In addition, these results supplement our previous findings in which greater FA was observed for the corpus callosum and tracts connecting frontal and parietal lobes (Luk et al., 2011). Furthermore, in a follow-up examination of resting state connectivity, we found that the frontoparietal control network was more strongly correlated with task-related executive function activity in bilinguals than in monolinguals (Grady et al., 2015). It has also been observed that bilinguals engage in a more distributed network of brain regions during language control tasks, relative to monolinguals consistent with the theoretical model proposed by Abutalebi & Green (2008) and results from a meta-analysis (Luk et al., 2012). The current results along with these previous studies support the notion that the structural maintenance of the neural connections amongst discrete frontal brain regions (as evidenced by greater frontal white matter and higher measures of white matter microstructure) may enable engagement of 1) a larger network of brain regions in lifelong bilinguals (i.e., the connections between nodes are more flexible or accessible via a larger number of routes); 2) the existing connections are stronger (i.e., greater myelination, preservation of the number of fibers, greater axonal caliber, etc.) by virtue of greater engagement of regions connected by these white matter pathways or 3) a combination of these structural changes may operate in parallel.

3.3 Potential mechanisms underlying changes in brain structure

For lifelong bilinguals, the experience of managing two languages is persistent and usually sustained by social circumstances, such as living in a multilingual community and maintaining connection with heritage and mainstream communities that do not share the same language. This experience-dependent situation potentially shapes brain functions and structures in order to cope with the demand of efficiently managing multiple language systems, including phonology, semantics, syntax and grammar (Costa & Sebastián-Gallés, 2014; Kroll & Bialystok, 2013). The increase in frontal white matter observed in the current study might reflect the continuous use of executive function mechanisms that regulate the retrieval of context-appropriate language and/or the suppression of the irrelevant

language. The association between the frontal lobe white matter and Stroop task performance provides an interesting linkage between bilingualism, executive function, and brain health in aging.

In the context of aging, it is reasonable that the prolonged experience of using two languages contributes to a mentally active lifestyle that facilitates the maintenance of brain functions and structures (Nyberg et al., 2012). Alternatively, bilingualism may contribute to cognitive reserve, or the ability for the brain to maximize cognitive performance in the face of age-related decline or disease (Stern, 2002). It is possible that both of these factors are at play: a more distributed functional and structural neural network may provide a "neural buffer" against regional declines associated with aging (brain reserve) and/or a more connected but distributed network as a result of lifelong bilingualism may impede or attenuate the impact of regional structural decline (compensation).

3.4 Limitations and future directions

One limitation of the current study is the relatively small sample size. It is possible that with a larger number of participants in each language group, the marginal difference in white matter volume observed in the temporal lobe would be significant. Similarly, while hippocampal volume and temporal pole grey matter thickness measures were numerically larger in bilinguals than in monolinguals, the group differences were not found to be significant as has been reported by others (Stein et al., 2012, Mårtensson, et al., 2012; Abutalebi et al., 2014). Furthermore, the cross-sectional nature of the study does not allow us to assume a causal relationship between lifelong bilingual experience and differences in brain structure or behavior in older adults. Follow-up longitudinal studies are necessary to determine the extent to which bilingual experience attenuates age-related changes in brain structure. Lastly, a definitive relationship between brain structure and neurocognitive function across the lifespan has yet to be established.

The current research extends well beyond the previous literature, in that we present data from both grey matter and white matter, and report data from neocortical structures and the hippocampus, whereas many previous papers examined either white matter properties alone (Luk et al., 2011; Schlegel et al., 2012) or grey matter properties in the neocortex (Stein et al., 2012, Mårtensson, et al., 2012; Abutalebi et al., 2014). Furthermore, we have examined the link between language experience, executive function and brain structures, as well as the relationship between age and brain structure. The current manuscript focused on T1-weighted images, which contain complementary information to diffusion-weighted images. We note that there is increasing evidence to suggest that differences in cortical thickness may also be related to underlying differences in white matter connectivity (Lerch et al., 2006; Raznahan et al., 2011). Future research should combine complementary, multimodal structural (e.g. volumetrics, cortical thickness, DTI) and functional imaging modalities as well as sensitive, age-appropriate behavioural assessments of cognitive function to fully understand the effects of lifelong bilingual experience.

4. Experimental procedure

4.1 Participants

Twenty-eight right-handed healthy older adults participated in the study. Fourteen participants were monolingual speakers of English (7 males and 7 females, *M* age = 70.6 years, *SD* = 3 years) and 14 had lifelong bilingual experience (6 males and 8 females; *M* age = 70.4 years, *SD* = 3.7 years). Participants provided informed consent and underwent a behavioural and a scanning session. The two groups had comparable demographic backgrounds and neuropsychological performance but different language experience (see Table 1). Two additional monolingual males were excluded from the analyses due to incidental findings on their MRIs. All the procedures were approved by the Research Ethics Board of the Baycrest Centre in Toronto, Canada. Monolingual older adults reported English to be their only communicating language, whereas the bilingual older adults reported that they had used both English and another alphabetic language regularly since childhood (before age 11). The monolingual and bilingual groups were not significantly different in terms of age and gender. All

participants were active community members, reported no known psychiatric or health issues that may affect neurological health, no experience of concussion, and no contraindication with MR scanning. Two-tailed t tests showed no statistical significant difference between the monolinguals and bilinguals in age, years of education, and weekly hours spent using a computer ($t < 2$, $p > 0.05$). An earlier neuroimaging investigation using this identical cohort was reported in Luk et al. (2011).

4.2 Neuropsychological tasks

Prior to MRI scanning session, participants underwent a 1 hour behavioural testing session consisting of a battery of neuropsychological tasks, including: Mini-Mental State Examination (short form, Folstein, Folstein, & McHugh, 1975), Shipley Institute of Living Scale—Vocabulary test (Zachary & Shipley, 1986), Verbal fluency test, the design fluency task from the Delis–Kaplan Executive Functions System (Delis, Kaplan, & Kramer, 2001), the Stroop task (Stroop, 1935), and Trail-making task (Reitan, 1958).

4.3 MRI acquisition

Approximately 2 weeks after the behavioural testing session, participants returned for the MRI scanning session. Anatomical (T1-weighted) data were acquired on a 3.0 tesla Siemens Trio scanner with a 12-channel head coil with the following parameters: 160 1-mm-thick oblique axial slices of 3D MPRAGE T1 images with TR = 2 s, TE = 2.63 ms, and FOV = 256 mm.

4.4 Volumetric analysis neocortical lobar volumes

The Minc Tool Kit (Montreal Neurological Institute, www.bic.mni.mcgill.ca/software/minc) was used for the structural analysis of the T1-weighted images based on a modified version of the ANIMAL algorithm (Collins, Holmes, Peters, & Evans, 1995). First, images were corrected for intensity inhomogeneities due to radio frequency (RF) field uniformity (Sled, Zijdenbos, & Evans, 1998) and slice-wise intensity normalization using the median of slice-wise intensity ratios (Zijdenbos, Forghani, & Evans, 2002). Next, spatial normalization was performed; the images underwent linear and nonlinear registration

(using the ANTs algorithm (Avants, Epstein, Grossman, & Gee, 2008)) to the ICBM-152 Atlas. Tissue classification priors predefined on this template were then projected back to the native space of each participant to enable a neural network-based classification of grey and white matter and cerebrospinal fluid (Tohka, Zijdenbos, & Evans, 2004). Similarly, a maximum probability atlas defining all of the major lobes of the brain (both grey and white matter) (Collins et al., 1995) was warped to match each participant and the intersection of this atlas with grey and white matter classifications to yield a lobe-wise segmentation mask (see Figure 1). The subject-specific whole brain segmentation mask, which was used for the creation of the lobar masks, was visually-inspected and hand-edited for all participants while blinded to language group status during this step. To ensure reliability of these hand edits, brain segmentation masks from seven randomly selected participants were hand-edited by a second rater, who was also blinded to group status, and the dice coefficient was computed, which produces an overlap measure between 0 and 1, where 0 signifies no overlap and 1 is a perfect match (Dice, 1945). The mean dice coefficient was found to be 0.97, which indicates that the hand-edits were reliable. This whole brain mask was then used to compute the volume for each lobar compartment (e.g., intracranial volume, left frontal grey matter volume, left frontal white matter volume, etc.). Lobar estimates were corrected/normalized by dividing each value by the corresponding intracranial volume (ICV) for each participant to account for differences in head size. ICV corresponds to the volume within the skull, containing brain and the surrounding tissues. All of the comparisons reported below were performed on the head-size normalized values (e.g., left frontal white matter $volume_{norm}$= left frontal white matter / ICV, frontal white matter $volume_{norm}$ = left frontal white matter + right frontal white / ICV).

4.5 Temporal lobe region of interest analysis

Cortical reconstruction and volumetric segmentation of the T1-weighted images was also performed with the FreeSurfer image analysis suite (Version 5; http://surfer.nmr.mgh.harvard.edu/). The technical details of these procedures are described in prior publications (Dale, Fischl, & Sereno, 1999; Dale & Sereno, 1993; Fischl & Dale, 2000; Fischl, Sereno, & Dale, 1999; Fischl, Sereno, Tootell, & Dale, 1999;

Fischl et al., 2002; Fischl, Salat, et al., 2004; Fischl, Liu, & Dale, 2001; Han et al., 2006; Jovicich et al., 2006; F Ségonne et al., 2004). Briefly, this processing involves the removal of non-brain tissue using a hybrid watershed/surface deformation procedure (Segonne et al., 2004), automated Talairach transformation, segmentation of the subcortical white matter and deep grey matter volumetric structures (including hippocampus, amygdala, caudate, putamen, and ventricles; Fischl et al., 2002; Fischl et al., 2004a) intensity normalization (Sled, Zijdenbos, & Evans, 1998), tessellation of the grey matter white matter boundary, automated topology correction (Fischl et al., 2001; Florent Ségonne, Pacheco, & Fischl, 2007), and surface deformation following intensity gradients to optimally place the grey/white and grey/cerebrospinal fluid borders at the location where the greatest shift in intensity defines the transition to the other tissue class (Dale et al., 1999; Dale and Sereno, 1993; Fischl and Dale, 2000). Once the cortical models were complete, a number of deformable procedures were performed for further data processing and analysis including surface inflation (Fischl et al., 1999a), registration to a spherical atlas, which utilized individual cortical folding patterns to match cortical geometry across subjects (Fischl et al., 1999b), parcellation of the cerebral cortex into units based on gyral and sulcal structure (Desikan et al., 2006; Fischl, van der Kouwe, et al., 2004), and creation of surface based data including maps of curvature. This method uses both intensity and continuity information from the entire three dimensional MR volume in segmentation and deformation procedures to produce representations of cortical thickness, calculated as the closest distance from the grey/white boundary to the grey/CSF boundary at each vertex on the tessellated surface (Fischl and Dale, 2000). All images were visually inspected prior to inclusion in the statistical analysis; no hand-edits were made to the FreeSurfer segmentations.

FreeSurfer provides volumetric measures for subcortical regions and both volumetric and cortical thickness measurements for neocortical regions. A region of interest approach, as opposed to a vertex-wise whole brain analysis, was used in order to maximize power to detect group differences in the current sample (Pardoe et al., 2013). Pardoe and colleagues have examined statistical power and

corrections for multiple comparisons of cortical thickness in a comparably aged sample. This research indicates that the current study is sufficiently powered for cortical thickness analyses of specific ROIs, but may be under-powered if strict multiple comparison correction for the large number of vertices included in whole-brain cortical thickness analyses (e.g., FDR corrections) are applied. We focused on the structures within the medial temporal lobe that are especially affected by aging and dementia (entorhinal cortex and hippocampus) as well as a region which is involved in language processing and was previously implicated in bilingual/monolingual comparisons within a similarly aged cohort (temporal pole; Abutalebi et al., 2014). The complex folding pattern of the hippocampus makes it difficult to accurately “unfold” or “flatten” this structure using currently available automated methods. Automated methods, such as FreeSurfer, typically unfold or flatten the cortex before computing the distance between the grey/white matter border and the grey/CSF border (Fischl and Dale, 2000; see Ekstrom et al., 2009 for a desciption of unfolding methods using high-resolution MRI). Instead, the most commonly used metric to assess structural differences in the hippocampus, at least based on standard resolution structural MRI (1x1x1mm voxel size), is volume. For this reason, we computed the volume and thickness of the entorhinal cortex and temporal pole, but only the volume of the hippocampus. The hippocampal volume of the oldest participant in our sample, a bilingual, was greater than 2.5 standard deviations below the group mean. We performed the volumetric analyses with and without this participant and the results were similar; the results reported above excluded this participant. As with the lobar volumes, temporal lobe regional volumes were divided by ICV to account for differences in head size. Cortical thickness was normalized using a global cortical thickness measure for each participant.

4.6 Statistical analysis

Statistical testing was performed using IBM SPSS Statistics (version 20) and SAS (version 9.3). Repeated measures ANOVAs were used to investigate potential differences in neocortical and hippocampal volumes as a function of language experience. Post-hoc comparisons were used to investigate effects underlying significant group, Group x Lobe or Group x Hemisphere interactions.

One-tailed tests were used for post-hoc comparisons as we predicted bilinguals would exhibit greater brain volumes than monolinguals. To compare the differences in the age-related trajectory between the two groups, mixed model regressions were run with group, age, and age by group interaction factors for each brain region. The significance of the age-related slopes for each group was assessed only in the presence of a significant Age x Group interaction.

Acknowledgments

This work was supported by the Canadian Institute of Health Research (MOP14036 to CG) and the Natural Sciences and Engineering Research Council of Canada (A2559 to EB). We would like to thank Joyce Chu and Erin Dickie for assistance with data analysis. We also thank Fergus Craik for helpful discussions regarding this investigation. Computations were performed on the gpc supercomputer at the SciNet HPC Consortium. SciNet is funded by: the Canada Foundation for Innovation under the auspices of Compute Canada; the Government of Ontario; Ontario Research Fund - Research Excellence; and the University of Toronto.

References

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*. doi:10.1016/j.neurobiolaging.2014.03.010

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., … Costa, A. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, *22*, 2076–2086. doi:10.1093/cercor/bhr287

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes*, *23*, 557–582. doi:10.1080/01690960801920602

Ahlskog, J. E., Geda, Y. E., Graff-Radford, N. R., & Petersen, R. C. (2011). Physical exercise as a preventive or disease-modifying treatment of dementia and brain aging. *Mayo Clinic Proceedings*, *86*, 876–884. doi:10.4065/mcp.2011.0252

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., … Kaul, S. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology*, *81*, 1938–1944. doi:10.1212/01.wnl.0000436620.33155.a4

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*, 26–41. doi:10.1016/j.media.2007.06.004

Bialystok, E., Craik, F. I. M., Binns, M. A., Ossher, L., & Freedman, M. (2014). Effects of bilingualism on the age of onset and progression of MCI and AD: Evidence from executive function tests. *Neuropsychology*, *28*, 290–304. doi:10.1037/neu0000023

Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, *45*, 459–464. doi:10.1016/j.neuropsychologia.2006.10.009

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*, 89–129. doi:10.1177/1529100610387084

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*, 290–303. doi:10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240–50. doi:10.1016/j.tics.2012.03.001

Bialystok, E., Craik, F. I. M., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *32*, 1341–1354. doi:10.1037/0278-7393.32.6.1341.

Bialystok, E., Poarch, G., Luo, L., & Craik, F. I. M. (2014). Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging*, 29(3), 696–705.

Bonner, M. F., & Price, a. R. (2013). Where is the anterior temporal lobe and what does it do? *Journal of Neuroscience*, 33(10), 4213–4215. doi:10.1523/JNEUROSCI.0041-13.2013

Brickman, A. M., Zimmerman, M. E., Paul, R. H., Grieve, S. M., Tate, D. F., Cohen, R. A., … Gordon, E. (2006). Regional white matter and neuropsychological functioning across the adult lifespan. *Biological Psychiatry*, *60*, 444–453. doi:10.1016/j.biopsych.2006.01.011

Cardenas, V. A., Chao, L. L., Studholme, C., Yaffe, K., Miller, B. L., Madison, C., … Weiner, M. W. (2011). Brain atrophy associated with baseline and longitudinal measures of cognition. *Neurobiology of Aging*, *32*(4), 572–580. doi:10.1016/j.neurobiolaging.2009.04.011

Colcombe, S. J., Erickson, K. I., Scalf, P. E., Kim, J. S., Prakash, R., McAuley, E., … Kramer, A. F. (2006). Aerobic exercise training increases brain volume in aging humans. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *61*, 1166–1170.

Collins, D. L., Holmes, C. J., Peters, T. M., & Evans, A. C. (1995). Automatic 3-D model-based neuroanatomical segmentation. *Human Brain Mapping*, *3*, 190–208. doi:10.1002/hbm.460030304

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

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, *9*, 179–194.

Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: A linear approach. *Journal of Cognitive Neuroscience*, *5*(2), 162–176.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX: Psychological Corporation.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. doi:10.1016/j.neuroimage.2006.01.021

Dice, L.R., 1945. Measures of the amount of ecologic association between species. *Ecology* 26, 297–302.

Ekstrom, A. D., Bazih, A. J., Suthana, N. a, Al-Hakim, R., Ogura, K., Zeineh, M., … Bookheimer, S. Y. (2009). Advances in high-resolution imaging and computational unfolding of the human hippocampus. *NeuroImage*, 47(1), 42–9. doi:10.1016/j.neuroimage.2009.03.017

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., … Kramer, A. F. (2011). Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 3017–22. doi:10.1073/pnas.1015950108

Ferreira, D., Molina, Y., Machado, A., Westman, E., Wahlund, L. O., Nieto, A., … Barroso, J. (2014). Cognitive decline is mediated by gray matter changes during middle age. *Neurobiology of Aging*, *35*, 1086–1094. doi:10.1016/j.neurobiolaging.2013.10.095.

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.

Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: Constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Transactions on Medical Imaging*, *20*(1), 70–80.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A. M. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*, 341–355. doi:10.1016/S0896-6273(02)00569-X.

Fischl, B., Salat, D. H., van der Kouwe, A. J. W., Makris, N., Ségonne, F., Quinn, B. T., & Dale, A. M. (2004). Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, *23*, S69–S84. doi:10.1016/j.neuroimage.2004.07.016.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–207.

Fischl, B., Sereno, M. I., Tootell, R. B. H., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*(4), 272–284.

Fischl, B., van der Kouwe, A. J. W., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., … Dale, A. M. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, *14*(1), 11–22. doi:10.1093/cercor/bhg087.

Fjell, A. M., Walhovd, K. B., Fennema-Notestine, C., McEvoy, L. K., Hagler, D. J., Holland, D., … Dale, A. M. (2009). One-year brain atrophy evident in healthy aging. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*, 15223–15231. doi:10.1523/JNEUROSCI.3252-09.2009

Fjell, A. M., Westlye, L. T., Grydeland, H., Amlien, I., Espeseth, T., Reinvang, I., … Walhovd, K. B. (2013). Critical ages in the life course of the adult brain: Nonlinear subcortical aging. *Neurobiology of Aging*, *34*(10), 2239–47. doi:10.1016/j.neurobiolaging.2013.04.006

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of pateitns for the clinician. *Journal of Psychiatric Research*, *12*, 189–198. doi:10.1016/0022-3956(75)90026-6

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience*, *23*, 9240–9245.

Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: what have we learned and where are we going? *Neuron*, *67*(5), 728–34. doi:10.1016/j.neuron.2010.08.040

Gold, B. T., Johnson, N. F., & Powell, D. K. (2013a). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia*, *51*(13), 2841–6. doi:10.1016/j.neuropsychologia.2013.09.037

Gold, B. T., Kim, C., Johnson, N. F., Kryscio, R. J., & Smith, C. D. (2013b). Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *The Journal of Neuroscience*, *33*, 387–396. doi:10.1523/JNEUROSCI.3837-12.2013.

Grady, C. L., Luk, G., Craik, F. I. M., & Bialystok, E. (2015) Brain network activity in monolingual and bilingual older adults. *Neuropsychologia*. 66,170-81. doi:10.1016/j.neuropsychologia.2014.10.042

Han, X., Jovicich, J., Salat, D., van der Kouwe, A. J. W., Quinn, B., Czanner, S., … Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*(1), 180–194. doi:10.1016/j.neuroimage.2006.02.051.

Hanna-Pladdy, B., & MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology*, *25*(3), 378–86. doi:10.1037/a0021895

Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*, *29*(4), 737–52. doi:10.1016/j.cger.2013.07.002

Head, D., Buckner, R. L., Shimony, J. S., Williams, L. E., Akbudak, E., Conturo, T. E., … Snyder, A. Z. (2004). Differential vulnerability of anterior white matter in nondemented aging with minimal acceleration in dementia of the Alzheimer type: Evidence from diffusion tensor imaging. *Cerebral Cortex*, *14*, 410–423. doi:10.1093/cercor/bhh003

Jack, C. R., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., … Kokmen, E. (1998). Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. *Neurology*, *51*, 993–999.

Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *NeuroImage*, *73*, 239–54. doi:10.1016/j.neuroimage.2012.06.081

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A. J. W., Gollub, R., … Dale, A. M. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, *30*(2), 436–443. doi:10.1016/j.neuroimage.2005.09.046.

Kennedy, K. M., & Raz, N. (2009). Aging white matter and cognition: differential effects of regional variations in diffusion properties on memory, executive functions, and speed. *Neuropsychologia*, *47*, 916–927. doi:10.1016/j.neuropsychologia.2009.01.001

Kramer, A. F., & Erickson, K. I. (2007). Capitalizing on cortical plasticity: influence of physical activity on cognition and brain function. *Trends in Cognitive Sciences*, *11*, 342–348. doi:10.1016/j.tics.2007.06.009

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*, 497–514. doi:10.1080/20445911.2013.799170 Understanding

Lenroot, R. K., Gogtay, N., Greenstein, D. K., Wells, E. M., Wallace, G. L., Clasen, L. S., … Giedd, J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *NeuroImage*, *36*(4), 1065–73. doi:10.1016/j.neuroimage.2007.03.053

Lenroot, R. K., Schmitt, J. E., Ordaz, S. J., Wallace, G. L., Neale, M. C., Lerch, J. P., … Giedd, J. N. (2009). Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. *Human Brain Mapping*, *30*(1), 163–74. doi:10.1002/hbm.20494

Lerch JP, Worsley K, Shaw WP, Greenstein DK, Lenroot RK, Giedd J, Evans AC (2006) Mapping anatomical correlations across cerebral cortex (MACACC) using cortical thickness from MRI. *Neuroimage* 31:993–1003.

Luk, G., Bialystok, E., Craik, F. I. M., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *The Journal of Neuroscience*, *31*, 16808–16813. doi:10.1523/JNEUROSCI.4563-11.2011

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes*, *27*, 1479–1488. doi:10.1080/01690965.2011.613209.

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, 63(1), 240–244. doi:10.1016/j.neuroimage.2012.06.043

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature*, *431*, 757. doi:10.1038/nature03016

Mueller, E. A., Moore, M. M., Kerr, D. C. R., Sexton, G., Camicioli, R. M., Howieson, D. B., … Kaye, J. A. (1998). Brain volume preserved in healthy elderly through the eleventh decade. *Neurology*, *51*, 1555–1562. doi:10.1212/WNL.51.6.1555

Nyberg, L., Lövdén, M., Riklund, K., Lindenberger, U., & Bäckman, L. (2012). Memory aging and brain maintenance. *Trends in Cognitive Sciences*, *16*, 292–305. doi:10.1016/j.tics.2012.04.005

Pardoe HR, Abbott DF, Jackson GD, The Alzheimer's Disease Neuroimaging (2013) Sample Size Estimates for Well-Powered Cross-Sectional Cortical Thickness Studies. *Human Brain Mapping* 34:1–19.

Petersen, R. C., Jack, C. R., Xu, Y.-C., Waring, S. C., O'Brien, P. C., Smith, G. E., … Kokmen, E. (2000). Memory and MRI-based hippocampal volumes in aging and AD. *Neurology*, *54*, 581. doi:10.1212/WNL.54.3.581

Raz, N., Gunning-Dixon, F., Head, D., Rodrigue, K. M., Williamson, A., & Acker, J. D. (2004). Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: replicability of regional differences in volume. *Neurobiology of Aging*, *25*(3), 377–396. doi:10.1016/S0197-4580(03)00118-0.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., … Acker, J. D. (2005). Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cerebral Cortex*, *15*, 1676–1689. doi:10.1093/cercor/bhi044

Raznahan A, Lerch JP, Lee N, Greenstein D, Wallace GL, Stockman M, Clasen L, Shaw PW, Giedd JN (2011) Patterns of coordinated anatomical change in human cortical development: a longitudinal neuroimaging study of maturational coupling. *Neuron* 72:873–884.

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*, 271–276. doi:10.2466/pms.1958.8.3.271

Resnick, S. M., Goldszal, A. F., Davatzikos, C., Golski, S., Kraut, M. A., Metter, E. J., … Zonderman, A. B. (2000). One-year age changes in MRI brain volumes in older adults. *Cerebral Cortex*, *10*, 464–472. doi:10.1093/cercor/10.5.464

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. *The Journal of Neuroscience*, *23*, 3295–3301.

Ross, L. a, McCoy, D., Wolk, D. a, Coslett, H. B., & Olson, I. R. (2010). Improved proper name recall by electrical stimulation of the anterior temporal lobes. *Neuropsychologia*, 48(12), 3671–4. doi:10.1016/j.neuropsychologia.2010.07.024

Salat, D. H., Greve, D. N., Pacheco, J. L., Quinn, B. T., Helmer, K. G., Buckner, R. L., & Fischl, B. (2009). Regional white matter volume differences in nondemented aging and Alzheimer's disease. *NeuroImage*, *44*, 1247–1258. doi:10.1016/j.neuroimage.2008.10.030

Salvatierra, J. L., & Rosselli, M. (2010). The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism*, *15*(1), 26–37. doi:10.1177/1367006910371021

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*, 1047–1055.

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), 1664–70. doi:10.1162/jocn_a_00240

Schroeder, S. R., & Marian, V. (2012). A bilingual advantage for episodic memory in older adults. *Journal of Cognitive Psychology (Hove, England)*, *24*, 591–601. doi:10.1080/20445911.2012.669367.

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, *22*(3), 1060–1075. doi:10.1016/j.neuroimage.2004.03.032

Ségonne, F., Pacheco, J., & Fischl, B. (2007). Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Transactions on Medical Imaging*, *26*(4), 518–529.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions on Medical Imaging*, *17*, 87–97. doi:10.1109/42.668698

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., … Dierks, T. (2012). Structural plasticity in the language system related to increased second language proficiency. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 48(4), 458–65. doi:10.1016/j.cortex.2010.10.007

Stein, M., Winkler, C., Kaiser, A., & Dierks, T. (2014). Structural brain changes related to bilingualism: does immersion make a difference? *Frontiers in Psychology*, 5(April), 1116. doi:10.3389/fpsyg.2014.01116

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*, 448–460. doi:10.1017.S1355617701020240

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *28*, 643–662. doi:10.1037/h0054651.

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Nagase, T., … Kawashima, R. (2012). Regional gray and white matter volume associated with Stroop interference: evidence from voxel-based morphometry. *NeuroImage*, 59(3), 2899–907. doi:10.1016/j.neuroimage.2011.09.064

Tohka, J., Zijdenbos, A., & Evans, A. (2004). Fast and robust parameter estimation for statistical partial volume models in brain MRI. *NeuroImage*, *23*, 84–97. doi:10.1016/j.neuroimage.2004.05.007

Tranel, D. (2009). The Left Temporal Pole Is Important for Retrieving Words for Unique Concrete Entities. *Aphasiology*, *23*(23), 867–884. doi:10.1080/02687030802586498

Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Dale, A. M., Eilertsen, D. E., … Fischl, B. (2005). Effects of age on volumes of cortex, white matter and subcortical structures. *Neurobiology of Aging*, *26*, 1261–1270. doi:10.1016/j.neurobiolaging.2005.05.020.

Walhovd, K. B., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., … Fjell, A. M. (2011). Consistent neuroanatomical age-related volume differences across multiple samples, *Neurobiology of Aging, 32*(5), 916–932. doi:10.1016/j.neurobiolaging.2009.05.013.

Wan, C. Y., & Schlaug, G. (2010). Music making as a tool for promoting brain plasticity across the life span. *The Neuroscientist*, *16*, 566–77. doi:10.1177/1073858410377805.

Wilson, R. S., Boyle, P. A., Yang, J., James, B. D., & Bennett, D. A. (in press). Early life instruction in foreign language and music and incidence of Mild Cognitive Impairment. *Neuropsychology*.

Zachary, R. A., & Shipley, W. C. (1986). *Shipley Institute of Living Scale: Revised Manual*. Los Angeles, CA: Western Psychology Services.

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*, 114–146. doi:10.1038/nn.3045

Zijdenbos, A. P., Forghani, R., & Evans, A. C. (2002). Automatic “pipeline” analysis of 3-D MRI data for clinical trials: Application to multiple sclerosi. *IEEE Transactions on Medical Imaging*, *21*, 1280–1291. doi:10.1109/TMI.2002.806283

Table 1.
Demographic information and neuropsychological performance

Measure	Monolingual (n=14)	Bilingual (n=13)
Age	70.6 (3.1)	69.9 (3.3)
MMSE – short form (out of 17)	16.9 (0.4)	16.9 (0.3)
Years of education	16.0 (2.8)	17.8 (2.0)
Shipley (English vocabulary)	89% (8%)	86% (13%)
Verbal fluency		
Letter fluency	13.6 (4.0)	15.4 (5.7)
Category fluency	18.4 (3.3)	20.1 (3.1)
Design fluency		
Baseline	10.9 (3.0)	10.8 (4.4)
Empty dots only	11.4 (2.6)	10.6 (3.2)
Switching	7.2 (3.0)	8.1 (2.6)
Stroop response time		
Baseline	23.7 (5.6)	23.6 (4.1)
Interference	37.8 (10.2)	36.4 (8.3)
Negative Priming	38.1 (13.2)	34.8 (6.8)
Habituation	30.3 (8.3)	30.7 (6.6)
Trail-making response time		
Numbers only	29.2 (10.8)	29.0 (10.4)
Letters only	30.0 (8.2)	28.3 (12.2)
Switching	70.9 (21.8)	65.7 (22.4)

Note. Abbreviations: MMSE = Mini-mental state examination

Table 2.
White matter volumes

	Bilinguals		Monolinguals		Effect of group		Age x Group interaction	
	M	*SD*	*M*	*SD*	*t*	*p*	*F*	*p*
Frontal White Matter	0.071	0.007	0.066	0.008	**1.67**	**0.05**	1.73	0.20
Temporal White Matter	0.045	0.005	0.046	0.004	1.37	0.09	2.86	0.10
Parietal White Matter	0.044	0.005	0.043	0.005	0.63	0.27	0.92	0.35
Occipital White Matter	0.021	0.002	0.021	0.003	-0.22	0.41	0.01	0.91

Note. Mean white matter $volume_{norm}$ and standard deviation are listed along with statistics associated with group comparisons.

Figure legends

Figure 1.

(A) T1-weighted image for a representative participant (upper) and T1-weighted images with the lobar masks overlaid (lower). Colour guide: Turquoise = Left frontal grey matter; light purple = left frontal white matter; red = right frontal grey matter; green = right frontal white matter; beige = left parietal grey matter; brown = left parietal white matter; dark blue = right parietal grey matter; light blue = right parietal white matter; dark brown = left temporal grey matter; olive = left temporal white matter; fuchsia = right temporal white matter; yellow = right temporal white matter; dark purple = right occipital grey matter; pink = right occipital white matter; light green = right cerebellar grey matter (not shown: left occipital grey and white matter, left cerebellar grey and white matter, and right cerebellar white matter). (B) White matter $volume_{norm}$ for the frontal, temporal, parietal, and occipital lobes, plotted separately for the two groups. Error bars represent the standard error.

Figure 2.

(A) T1-weighted image for a representative participant with left and right hippocampal ROIs overlaid in red. (B) Hippocampal $volume_{norm}$ for the left and right hippocampus, plotted separately for the two groups. Error bars represent the standard error.

Figure 3.

(A) Surface model created using FreeSurfer which depicts the left temporal pole region of interest overlaid in yellow on the template (“fsaverage”) brain. (B) Scatterplots depict the relationship between age and temporal pole grey matter (cortical $thickness_{norm}$) for the monolingual (left) and bilingual (right) groups separately. A significant AgeXGroup interaction was observed in this region.

Figure 4.

Relationship between frontal lobe white matter and Stroop task performance (interference condition – baseline) collapsed across language groups (A) and plotted for each group separately (B). Negative relationship indicates that better performance is associated with greater white matter volume in the frontal lobe.

Figure 5.

Schematic illustration of grey and white matter volumes in monolinguals and bilinguals across the lifespan. The current study examined participants ranging from 65 to 76 years. (A) Language learning

has been associated with increased temporal pole grey matter in young adults (Stein et al., 2012). Cortical thickness/volume in the temporal pole demonstrated a significant interaction with age and group in the current study as well as in Abutalebi et al. (2014), suggesting that age related declines are more pronounced in this region in monolinguals compared to bilinguals. (B) Hypothetical white matter curves depicting a main effect of group on white matter volume as observed in the current sample. Little is known about the difference in white matter volume between groups across the lifespan.

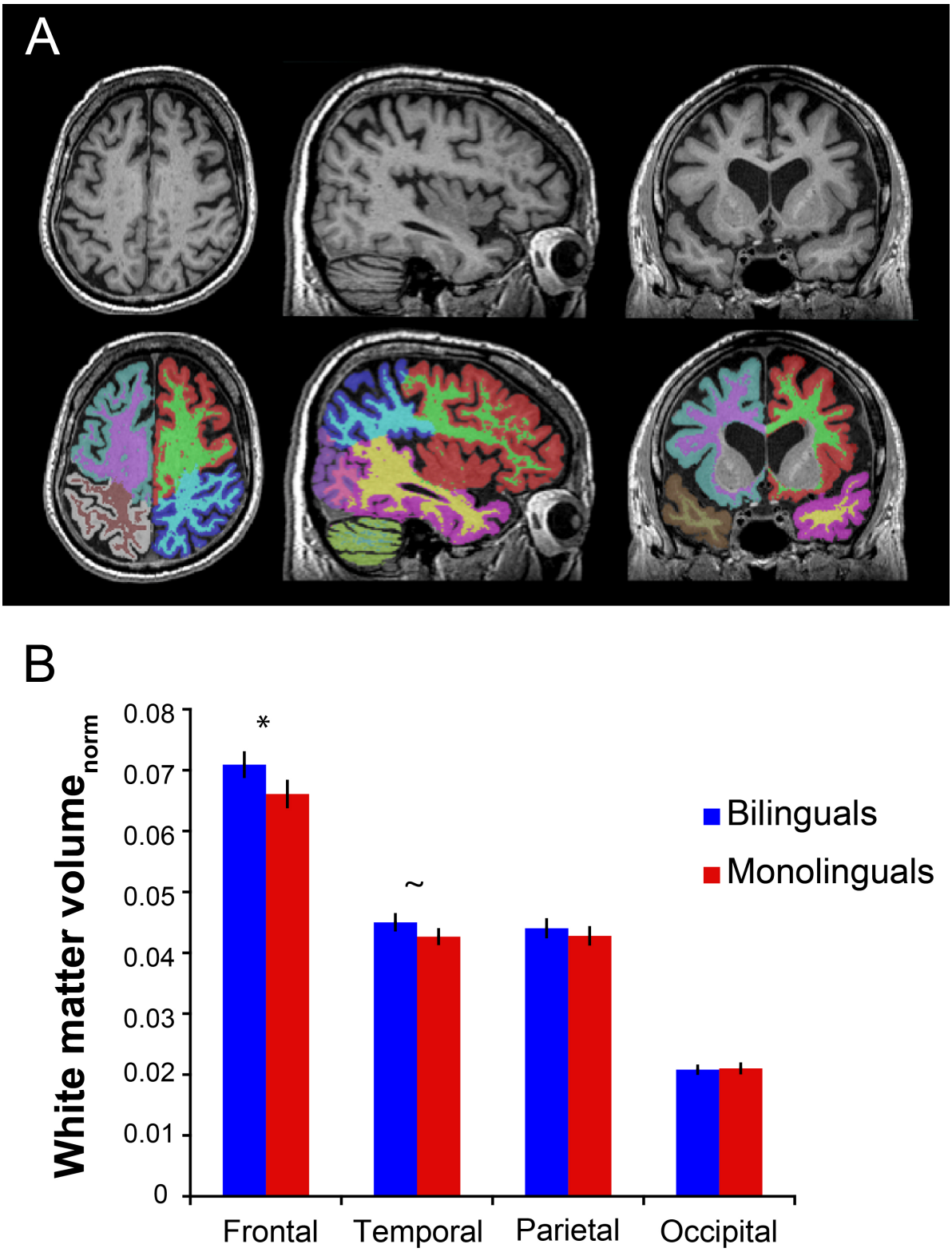
A
B
White matter volume norm
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
*
~
Bilinguals
Monolinguals
Frontal
Temporal
Parietal
Occipital

Fig 1.

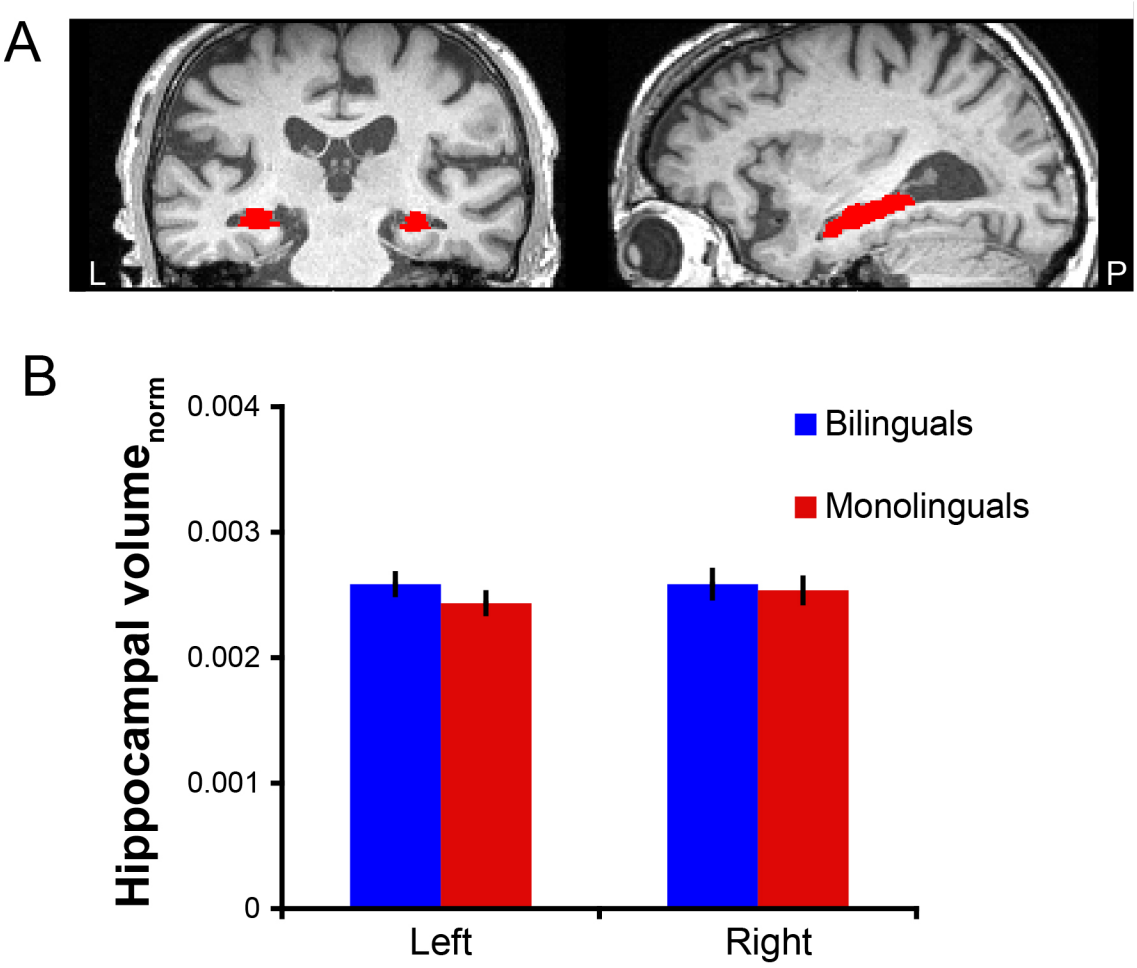
A
L
P
B
Hippocampal volume norm
0.004
0.003
0.002
0.001
0
Bilinguals
Monolinguals
Left
Right

Fig 2.

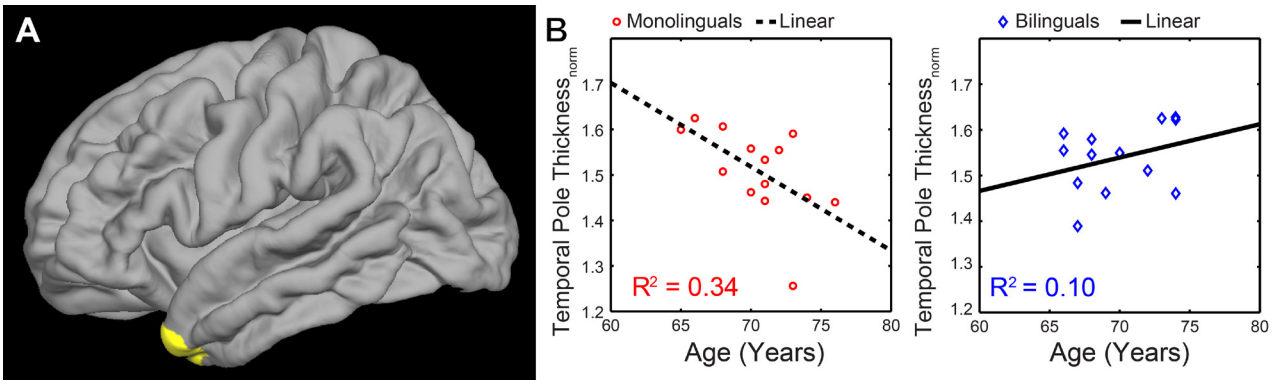
A
B
Monolinguals
Linear
Temporal Pole Thickness$_{norm}$
Age (Years)
R^2 = 0.34
Bilinguals
Linear
Temporal Pole Thickness$_{norm}$
Age (Years)
R^2 = 0.10

Fig 3.

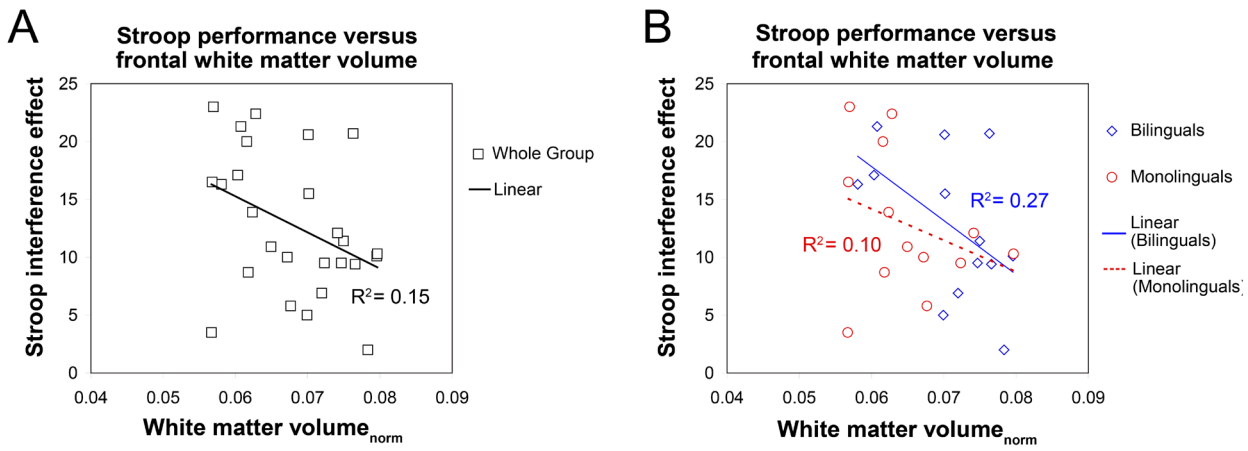
A
Stroop performance versus frontal white matter volume
Stroop interference effect
White matter volume norm
R² = 0.15
Whole Group
Linear
B
Stroop performance versus frontal white matter volume
Stroop interference effect
White matter volume norm
R² = 0.27
R² = 0.10
Bilinguals
Monolinguals
Linear (Bilinguals)
Linear (Monolinguals)

Fig 4.

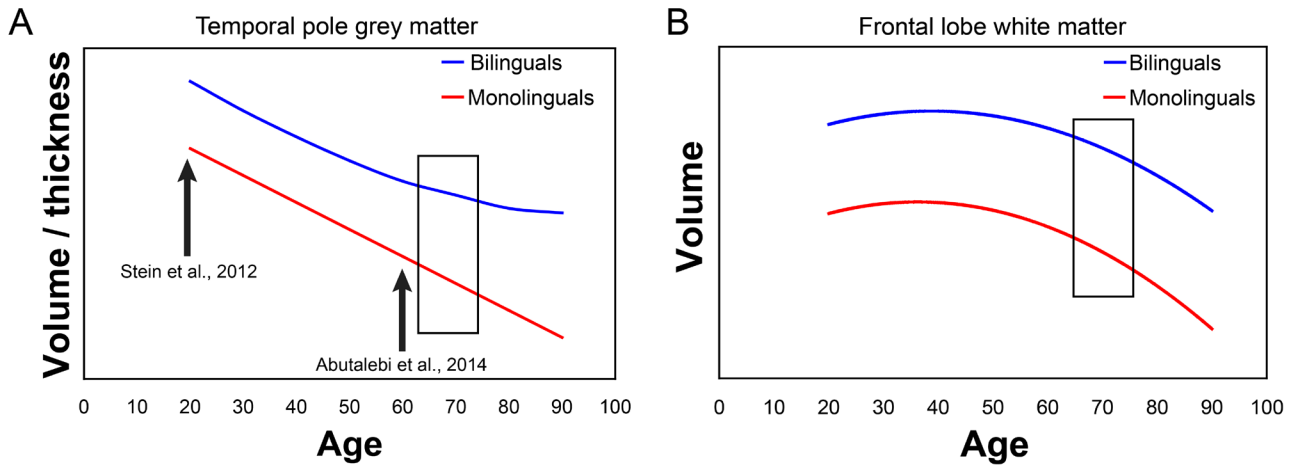
A
Temporal pole grey matter
Bilinguals
Monolinguals
Volume / thickness
Stein et al., 2012
Abutalebi et al., 2014
0 10 20 30 40 50 60 70 80 90 100
Age
B
Frontal lobe white matter
Bilinguals
Monolinguals
Volume
0 10 20 30 40 50 60 70 80 90 100
Age

Fig 5.