

Bilingualism Provides a Neural Reserve for Aging Populations

Jubin Abutalebi^{1,2}, *Lucia Guidi*^{2,3}, *Virginia Borsa*², *Matteo Canini*^{2,4}, *Pasquale A. Della Rosa*⁴,
*Ben A. Parris*⁵, & *Brendan S. Weekes*¹

¹ Department of Speech and Hearing Sciences, University of Hong Kong, Hong Kong

² Centre of Cognitive Neurosciences, San Raffaele University & San Raffaele Scientific Institute,
Milano, Italy

³ Istituto Universitario degli Studi Superiori – IUSS, Pavia, Italy

⁴ IBFM-CNR (National Research Council), Milan, Italy

⁵ Psychology Department, Bournemouth University, United Kingdom

Send correspondence to:

Dr. Jubin Abutalebi

Faculty of Psychology

University Vita-Salute San Raffaele,

Via Olgettina 58

20132 Milan, Italy

tel.: 0039 0226434888

email: abutalebi.jubin@hsr.it

Abstract

It has been postulated that bilingualism may act as a cognitive reserve and recent behavioral evidence shows that bilinguals are diagnosed with dementia about 4-5 years later compared to monolinguals. In the present study, we investigated the neural basis of these putative protective effects in a group of aging bilinguals as compared a matched monolingual control group. For this purpose, participants completed the Erikson Flanker task and their performance was correlated to grey matter (GM) volume in order to investigate if cognitive performance predicts GM volume specifically in areas affected by aging. We performed an ex-Gaussian analysis on the resulting RTs and report that aging bilinguals performed better than aging monolinguals on the Flanker task. Bilingualism was overall associated with increased GM in the ACC Task performance correlated with decreased GM in the DLPFC only in monolinguals. Taken together, these neural regions might underlie the benefits of bilingualism and act as a neural reserve that protects against the cognitive decline that occurs during aging..

Keywords: bilingualism, cognitive reserve, neural reserve, ex-Gaussian analysis, voxel based morphometry,

Abbreviations: DLPFC (dorsolateral prefrontal cortex), ACC (anterior cingulate cortex), RTs (reaction times)

Introduction

Cognitive decline is characterized by a slow and progressive decline of memory and cognitive abilities resulting from pathological conditions that cause brain cell death in the aging population. Although treatments are still lacking, studies show that the onset of cognitive decline and dementia can be significantly delayed by intellectual and life-style factors, including education, occupation level and leisure activities. These individual factors contribute to 'cognitive reserve' in seniors. Higher levels of cognitive reserve are associated with a reduced risk of developing cognitive decline and a lower rate of memory decline in normal aging (Stern et al., 2006). Stern et al. (2009) suggested that life experiences contribute to individual differences at the brain level, i.e., providing a neural basis for the concept of cognitive reserve (to which we may refer to as 'neural reserve'). The benefit of intellectual and life-style cognitive reserve is potentially considerable, given the socio-economic and affective impact of cognitive decline in typical aging. Delaying the onset of cognitive decline can lead to substantial economic benefits (Bialystok et al., 2004).

The concept of cognitive reserve is based on multiple observations that particularly intelligent people with a high level of education seem to fare better in the face of cognitive decline, maintaining cognitive function for a longer period of time compared to people with lower levels of education. Epidemiological studies (Gatz et al., 2006; Garibotto et al., 2008; for a review see Valenzuela & Sachdev, 2006) have demonstrated that low levels of education are strongly correlated with a higher risk of cognitive decline. Aging-induced decline in cognitive abilities is usually associated with structural and functional changes in the brain. For instance, progressive loss of grey matter volume in anterior regions is typically observed (Jernigan et al., 2001; Raz et al., 2005) as is progressive disruption of white matter between anterior and posterior cortical regions (Pfefferbaum et al., 2005; Head et al., 2004; Madden et al., 2009; Gunning-Dixon et al., 2009).

It has been postulated that bilingualism may act as a cognitive reserve. This notion is based upon evidence that executive functions may be better 'trained' in bilinguals as compared to monolinguals (Bialystok et al., 2004, 2005, 2006; Bialystok & Craik, 2010b). Indeed, there is a body of evidence showing that the constant juggling of two (or more) languages in multi- and bilingual speakers tunes their conflict resolution capacity, which in turn, transfers to non-linguistic behavioral and cognitive functions. Specifically, domain-general executive functions might benefit from brain plasticity induced by bilingualism, which then allows greater resistance to cognitive decline (Abutalebi et al., 2012; Zou et al., 2012). Brain areas underlying executive control, such as the pre-SMA/ACC and the DLPFC regions are more highly stimulated in bilingual speakers and this may result in greater cognitive reserve that compensates for the brain atrophy found in normal aging. Indeed, behavioral studies in elderly bilinguals support this notion (Bialystok et al., 2007; Bialystok et al., 2009). For instance, Bialystok and colleagues (2004) investigated the effect of aging on executive processes in a relatively large group of monolingual and bilingual speakers. Participants were divided into five different age group ranges (30-39; 40-49; 50-59; 60-69; and 70-79) and all performed an attentional task (i.e., Simon Task). The study found that bilinguals were faster in resolving the cognitive conflicts on this specific task. Moreover, the advantage for bilinguals was more evident with increasing age. Both monolingual and bilingual speakers show effects of slowing on task performance with age. However, the difference between the two groups widened with age. Several behavioral studies report a 'bilingual advantage' on executive functions in almost every age group (infants, children and adults) (Bialystok & Craik, 2010b; Kovács & Mehler, 2009). It should be noted that this evidence is not universally accepted (see Paap & Greenberg, 2013) especially for studies focusing on children and young adults. Following Valian (2015) children and young adults engage in many cognitively challenging activities that may be at least equivalent to the cognitive challenges provided by bilingualism. However, old adults tend to have fewer cognitively

enriching experiences than younger adults, and hence, any putative advantage provided by bilingualism could be more prominent (Valian, 2015). A further hypothesis is that eventual cognitive advantages increase with experience during aging, suggesting age dependent development of cognitive reserve (Hilchey & Klein, 2011).

As to potential neural repercussions, it is well documented that bilingualism induces beneficial experience-related structural changes in terms of increased grey and white matter densities when compared to monolingual speakers. Bilingualism induces structural changes in several brain areas including the left inferior parietal lobule (Mechelli et al., 2004; Della Rosa et al., 2013), the anterior cingulate cortex (ACC) (Abutalebi et al., 2012), and in subcortical structures such as the left caudate (Zou et al., 2012), and putamen (Abutalebi et al., 2013a). These areas are part of the executive control network and this may explain why bilinguals usually have a cognitive advantage in executive control tasks. Interestingly and specifically related to aging, Luk et al. (2011) report that aging bilinguals have globally increased white matter when compared to their monolingual peers. Likewise, Abutalebi et al. (2014) also showed that aging bilinguals have increased grey matter densities over the temporal poles and orbito-frontal cortex. These neural data fit relatively well with the recent evidence that bilingualism delays the onset of dementia (Bialystok et al., 2012). For instance, bilingual participants are diagnosed with dementia of the Alzheimer's Type (DAT) about 4-5 years later compared to monolinguals (Schweizer et al., 2012, Craik et al., 2010). These findings were recently confirmed and generalized to other types of dementia (such as vascular dementia) in a large study conducted in India (Alladi et al., 2013).

The main aim of the present study is to investigate the neural basis of the bilingual advantage in a group of aging bilinguals as compared to a matched monolingual control group. For this purpose, subjects carried out an attentional control task (i.e., the Flanker task) and their

performance - measured in reaction times (RTs) - was correlated to brain structure. The Flanker task captures cognitive conflict resolution involving attentional control and inhibition processes (Fan et al., 2002). Horizontal black lines with arrowheads pointing to left or right were presented (see Figure 1a for examples) and participants were instructed to indicate the direction of the central target arrow by pressing a dedicated response button as fast as possible. The central target is presented with congruent, incongruent or neutral flankers. The conflict effect is calculated by subtracting response times to congruent trials from those to incongruent trials (Fan et al., 2002). The Flanker task is widely used in studies of the bilingual cognitive advantage (Abutalebi et al., 2012; Calabria et al. 2011; Costa et al., 2008).

(INSERT FIGURE 1 OVER HERE)

The ex-Gaussian distribution analysis

The distribution of RTs in attention tasks is usually positively skewed. However classical measures of central tendency (Gaussian component) neglect this variability and therefore could potentially incur the risk of masking information about the real trend of the dispersion and unique characteristics of the populations sampled, possibly leading to a Type 2 error. To provide a better fit of the RT distribution, the Ex-Gaussian function involves a mathematical convolution of a normal (Gaussian) and an exponential distribution, producing the following 3 parameters: μ (Mu), reflecting the mean of the Gaussian component of the distribution; σ (Sigma), the standard deviation of the Gaussian distribution, and τ (Tau), the mean and the standard deviation of the exponential component. The sum of μ and τ yields the total mean of the distribution, whereas adding the square of the standard deviation of σ and τ draws its variance. The leading edge of the distribution is reflected in μ , whereas the skewness of the distribution is better reflected in τ .

Application of the ex-Gaussian function has been successfully employed to compare differences between groups, such as healthy elderly subjects when compared with individuals who have dementia (Spieler et al., 1996; Verhaeghen & Hoyer, 2007). Recently, Tse et al. (2010) showed that healthy aging had clear effects on both μ and τ in a set of attentional control tasks, whereas early stage DAT only had an additional effect on τ . Balota et al. (2010) demonstrated that the slow tail of the RT distribution in the Stroop task may also be useful in predicting divergence from a cognitive normal state to early stage DAT across a 12 year longitudinal follow up study. Jackson et al. (2012) showed a significant association between measures of intra-individual variability (IIV) reflected in the RTs of an attentional control task, the parameter τ , reflecting an exaggeration of the slow tail of the RTs, and white matter volume in an aging population and in early-stage DAT individuals, concluding that IIV and τ are sensitive to breakdowns in executive control processes in healthy and pathological aging. In sort, applying the Ex-Gaussian function to reaction times allows the splitting of a normal distribution into sub-components (μ and τ) that appear to be independently useful in predicting cognitive decline. It has been argued that between-group differences in the normal component μ are more likely to reflect differences in more “automatic” processes across the whole task, whereas between-group differences in the exponential component τ have been related to more “controlled” processes across the task (such as inhibition) (for a critical review see: Matzke & Wagenmakers, 2009).

Of note, and related to bilingualism, Calabria et al. (2011) used ex-Gaussian analysis to investigate potential cognitive advantages in a large group of young bilinguals given the Flanker task. They found that bilinguals perform faster on both components μ and τ , for incongruent and congruent trials and also had a smaller conflict effect. Calabria et al. also performed a correlation analysis between the magnitude of μ and τ and found that μ and τ were positively correlated for monolinguals whereas in the bilingual group they were not, suggesting that these parameters are

independent in bilinguals but not in monolinguals. They argued that the ex-Gaussian analysis was informative for investigating the bilingual advantage in young adults. No study has yet investigated the bilingual advantage in elderly bilinguals using ex-Gaussian analysis.

As aforementioned, in the present study we employed the Ex-Gaussian analysis approach in order to analyze RTs measured from the Flanker task in aging bilinguals. Participants also underwent structural MRI scanning to obtain grey matter volume indexes to permit correlations of these results with behavioral performance. Following the results from Calabria et al. (2011), we predicted that bilinguals would perform faster on both components μ and τ , for incongruent and congruent trials and would also show a smaller conflict effect. Furthermore, we predicted differential correlations between components and specific neural structures for cognitive processing in the bilingual and monolingual groups.

Materials and Methods

Subjects

30 healthy bilinguals from Hong Kong (16 Cantonese-English, 14 Cantonese-Mandarin; 13 males; mean age 63.2; standard deviation [SD] \pm 5.86; age range 55-75) and 30 healthy Italian monolinguals (14 males; mean age 61.85; standard deviation [SD] \pm 6.71; age range 49-75) participated in this study. The monolingual control group was chosen from the aging population in Milan because of the relative difficulty of finding real monolinguals in Hong Kong. Usually, the few monolinguals in Hong Kong and surrounding areas have a much lower SES (socio-economic status) and education than the local bilingual population. Milan and Hong Kong we are in many ways similar in terms of lifestyles. Both cities are economic hubs with a very comparable “rhythm of life”. Moreover, the aging population in Milan are not exposed to a

second language unlike in Hong Kong where even monolinguals are potentially exposed to second languages such as English and Mandarin.

Exclusion criteria were history of neurological/psychiatric disease or head injury and a minimum score on the MMSE of 27. We also applied exclusion criteria based on accuracy performance on the flanker task in order to increase the reliability of the QML method used in the estimation of ex-Gaussian parameters that maximize the likelihood of the sample data (see below for detailed description of the accuracy-based exclusion criteria). Applying the MMSE criteria, 7 out of 30 bilinguals and 2 out of 30 monolinguals were excluded. Following exclusion based on accuracy performance and MMSE exclusion criteria, 21 bilinguals and 24 Monolinguals remained. In order to ensure comparability of socio-demographic data between the bilingual and monolingual group while ensuring a sufficient and equal sample size, we further excluded 2 bilingual subjects and 5 monolingual subjects based on accuracy percentage tending towards the upper bound of the confidence interval for each group, but inside the CI (see below for a detailed explanation).

The final number of participants included in the present study was 19 for the bilingual group (11 Cantonese-English, 8 Cantonese-Mandarin) (8 males; mean age 61.68; standard deviation [SD] \pm 5.34; age range 55-75) and 19 for the monolingual group (9 males; mean age 60.93; standard deviation [SD] \pm 5.81; age range 49-75). Subjects had a comparable level of education (years of education for bilinguals $m=$ 13.76, [SD] \pm 3.98; monolinguals $m=$ 13.16, [SD] \pm 4.86), MMSE (raw score for bilinguals $M=$ 28.63, [SD] \pm 0.9 and monolinguals $M=$ 28.95, [SD] \pm 1.03; corrected score for bilinguals $M=$ 27.69, [SD] \pm 1 and monolinguals $M=$ 28.31, [SD] \pm 1.17), accuracy in the performance of the Flanker task (bilinguals $M=$ 99.21%, [SD] \pm 0.01; monolinguals $M=$ 99.23%, [SD] \pm 0.01) and Socio Economic Status (SES)¹

¹ The Socio-Economic Status (SES) of our sample was assessed by administering a questionnaire (SES, MacArthur Foundation Network, <http://www.macses.ucsf.edu/research/socialenviron/sociodemographic.Php>) aimed at investigating i) the subject's self-perceived social position, with respect to his/her local community and his/her country, on a 10 points rating scale; ii) the degree of education (i.e. the number of years of formal education) and iii) the personal and overall family income of the last 12 months. For this last dimension we adopted different salary ranges for the Hong Kong and Milan societies in order to create meaningful estimates of lower, middle and upper classes for both societies.

(bilinguals $M=22.18$, $[SD] \pm 5.39$; monolinguals $M= 22.55$, $[SD] \pm 4.12$). Independent t -tests showed that the groups did not differ significantly for any of the matching criteria, e.g. age ($p = .680$), MMSE raw score ($p= .416$) and corrected score ($p = .099$), Education ($p = .626$) and SES ($p = .813$) (see Table 1). The Human Research Ethics Committee at the University of Hong Kong approved the study. Informed consent was obtained from all participants in the study.

(INSERT TABLE 1 OVER HERE)

Second Language Competence Assessment

Bilingual participants were tested in L1 (first language) and L2 (second language) for the following variables (see Table 2): Age of acquisition of the second language (AoA); a picture naming task in L1 and L2 providing a quantitative measure of proficiency (30 items per language selected from the Snodgrass and Vanderwart revised battery (1980); a translation task from L1 to L2 and vice versa; a self-report questionnaire providing information about the amount of exposure to the second language (L2 Exposure) and usage of L2 over time (see for details Abutalebi et al., 2007).

(INSERT TABLE 2 OVER HERE)

Behavioral Study

Experimental procedures

All participants (bilinguals and monolinguals) performed the Flanker Task (Fan et al., 2002). The task was presented on a computer screen. Event presentation consisted of a fixation point appearing at the center of the screen for 400ms, followed by a row of five horizontal black lines

with arrowheads pointing to left or right for 1700ms (see Figure 1). Participants were instructed to indicate the direction of the central target arrow by pressing a dedicated response button as fast as possible. The central target arrow was presented with congruent, incongruent or neutral flankers (see Figure 1 for samples). The task was presented with different cues (no cue, center cue, double cue and spatial cue) (Fan et al., 2005) allowing the eventual investigation of the alerting and orienting components of attentional control (Posner & Petersen, 1990). Since our focus here was on the conflict effect, the alerting and orienting components were not analyzed. The experiment consisted of 2 runs with 96 items per run (33% congruent, 33% incongruent and 33% neutral) presented randomly. The conflict effect was calculated as the value resulting from subtracting responses of congruent trials from incongruent trials presented across the 2 runs.

Ex-Gaussian Method and Statistical Analysis

In order to perform the Ex-Gaussian analysis, data from the first and second runs of the Flanker task were plotted together. The original data were sorted per condition (Congruent, Incongruent, Neutral) and per group (monolinguals, bilinguals) separately in order to first carry out an analysis on simple reaction time (RTs) data. Subsequently, we carried out the Ex-Gaussian analysis with the quantile maximum likelihood (QML) procedure as implemented in version 2.18 of the QMPE software (Heathcote, et al., 2004). The Ex-Gaussian function is a mathematical convolution of the normal (Gaussian) and the exponential distribution (Heathcote et al., 1991) which produces estimates of the μ , σ , and τ parameters. Mathematically, the Ex-Gaussian probability function is expressed as $f(x|\mu, \sigma, \tau) = \frac{1}{\tau} \exp\left(\frac{\mu}{\tau} + \frac{\sigma^2}{2\tau^2} - \frac{x}{\tau}\right) \Phi\left(\frac{x-\mu-\sigma^2/\tau}{\sigma}\right)$ (Luce, 1986), where the exponential function is multiplied by the value of the cumulative density of the Gaussian function (Φ). The resulting Ex-Gaussian parameters are, respectively, i) μ corresponding to the mean of the Gaussian component of the distribution; ii) σ reflecting the standard deviation of the Gaussian distribution and iii) τ , the mean and the standard deviation of the exponential component. The leading edge of

the distribution is reflected in μ , whereas the skewness of the distribution is better reflected in τ (Heathcote, 1996). The QML is a method that combines the robustness of quantiles and the efficiency of maximum likelihood estimation, determining the parameters that maximize the probability of the set of data (for details, see Heathcote et al., 2002). We calculated the Ex-Gaussian parameters for each participant and for each type of trial (Congruent, Incongruent and Neutral). For each condition (congruent, incongruent and neutral) we had 64 observations per subject resulting in 2432 observations in total. From this amount we excluded those trials that were classified as “incorrect” (when the subjects failed to select the correct response) and “miss” (when the subjects did not make a response). RT data were not trimmed (i.e. removing outliers above 2\3 SDs) in order to avoid a potential loss of information at the level of ex-Gaussian analysis. Statistical analyses were carried out on μ and τ parameters. Independent sample t-tests were performed in order to test differences between groups on μ and τ parameters in the congruent and incongruent conditions. Furthermore, correlation analyses were carried out between age, τ and μ calculated for congruent and incongruent trials and for the difference between incongruent and congruent trials (Conflict Effect).

As mentioned above, we applied exclusion criteria based on accuracy performance on the flanker task. We, hence, took into account knowledge of response accuracy in order to better interpret potential differences between monolinguals and bilinguals in terms of different cognitive processes underlying conflict resolution measured as a function of response speed (i.e. ex-Gaussian μ and τ parameters) while controlling for accuracy of the original sample ($n=30$ subjects). This was done in order to minimize differences attributable to participant ability, response caution or to task difficulty, which can be affected by aging.

The first criterion was global accuracy performance on the Flanker task. Mean accuracy percentage scores (i.e. calculated as the average of the two Flanker Task sessions) were 98.33% (95% confidence interval [CI], 92.27%-104.40%) for the bilingual group and 93.77% (95%

confidence interval [CI], 81.90%-105.64%) for the monolingual group. One bilingual and two monolinguals were excluded due to falling below the lower CI boundary.

The second criterion was the difference in accuracy between the congruent and incongruent trials (i.e. collapsed across both Flanker Task sessions). Mean accuracy differences (i.e. Congruent minus Incongruent) expressed as percentages were 2.71% (95% confidence interval [CI], -7.56%-12.97%) for the bilingual group and 6.20% (95% confidence interval [CI], -21.48%-33.88%) for the monolingual group. One bilingual was excluded due to falling above the upper CI boundary. The third criterion was the difference in accuracy between the congruent and incongruent trials between the two Flanker Task sessions (i.e. accuracy session effect). Mean accuracy differences (i.e. Congruent minus Incongruent Session 1 - Congruent minus Incongruent Session 2) expressed as percentages were 1.74% (95% confidence interval [CI], -7.21%-10.68%) for bilinguals and 0.89% (95% confidence interval [CI], -8.06%-9.84%). One monolingual subject was excluded, due to falling above the upper CI boundary. Mean accuracy percentage values and confidence intervals for each criteria are shown for the bilingual and monolingual group separately as well as for the total sample ($n=38$ subjects) in Table 3.

(INSERT TABLE 3 OVER HERE)

Behavioral Results

ANOVA on RTs for Bilinguals and Monolinguals

A 2 x 2 ANOVA was carried out on RT data using Group (bilingual-monolingual) as between subjects factor and Congruency (Cong vs Incong) as within subjects factor. The ANOVA showed a main effect of Group [$F(1,36)=5,851$] ($p=0.021$; $\eta^2_p=0.14$) indicating that bilinguals (mean=669.64 ms; st.dev.= 20.55) were faster than monolinguals (mean=739,93 ms; st.dev.=20.55), and also a main effect of Congruency [$F(1,36)=284,29$] ($p<0.001$; $\eta^2_p=0,89$) indicating that for both groups responses on congruent trials (mean=648,79 ms; st.dev.=13,13)

were faster than those on incongruent trials (mean=760.74 ms; st.dev.=16.5). No Group x Congruency interaction was found [$F(1,36)=1.22$] ($p = 0.277$; $\eta^2_p = 0.033$). Post Hoc multiple comparisons (Bonferroni adjusted) revealed that the bilingual advantage over monolinguals was present for both congruent (mean RTs_BIL= 617.29 ms, st.dev=18.56; mean RTs_MONO= 680.29 ms, st.dev= 18,57) ($p=0.022$) and incongruent trials (mean RTs_BIL= 721.92ms, st.dev=23.33; mean RTs_MONO= 799.57 ms, st.dev=23.33) ($p=0.024$).

Analysis of Ex-Gaussian parameters: Bilinguals vs Monolinguals

Independent sample t -tests revealed differences between monolinguals and bilinguals in the τ component calculated for incongruent trials (τ -incong) and the μ component calculated for congruent trials (μ -cong) (Figure 2). Monolinguals exhibited significantly greater values of τ -incong than bilinguals ($t = -2,219$, $df=36$, $p = .033$) and significantly greater values of μ -cong compared to bilinguals ($t = -2,28$, $df=36$, $p = .029$). Furthermore, a trend towards significance was found for μ -incong trials ($p=.069$) with monolinguals having larger values. No significant differences between the groups emerged for the conflict effect in μ ($p = .485$) or in τ ($p = .535$). To summarize, the independent sample t -tests revealed a significant advantage for bilinguals compared to monolinguals for the μ component on congruent trials and for the τ component on incongruent trials.

Correlation analysis

Correlation analyses were also run in order to investigate any possible correlation between the μ and τ components calculated for congruent trials (μ -cong, τ -cong), incongruent trials (μ -incong, τ -incong), the conflict effect (μ -CE, τ -CE) and the age variable, given that our main interest was to assess the factor age in the two groups of aging subjects. Interestingly, only in monolinguals was a significant positive correlation found between age and μ -incong ($r=0.541$; $p = .017$) and between age and μ -cong ($r = 0.496$, $p = .031$). We also found a significant negative correlation between age

and τ -CE ($r = -0.573$, $p = .010$). Furthermore, for both groups a positive correlation was found between τ -cong and μ -incong (bilingual: $r = 0.476$, $p = .040$; monolingual: $r = 0.522$, $p = .022$).

(INSERT FIGURE 2 OVER HERE)

Structural Neuroimaging Study

Data Acquisition: Bilinguals

High-resolution T1 structural images were acquired for the 30 bilinguals at the 3T MRI center of the University of Hong Kong using a 3T Achieva Philips MR scanner (Philips Medical Systems, Best, the Netherlands). An axial high-resolution structural MRI scan was acquired for each subject (magnetization prepared rapid gradient echo, 150 slice T1-weighted image, repetition time = 8.03 ms, echo time = 4.1 ms; flip angle = 8°, field of view = 250x250, matrix = 256, acquisition time (TA) = 9.35 min, Mode = 3D fast-field echo (3DFFE), sense factor = 1, number of signal averages = 1, resolution = 1x1x1).

Data Acquisition: Monolinguals

High-resolution T1 structural images were acquired for the 30 monolingual participants at the CERMAC center at University San Raffaele in Milan (Italy). The same scanner model (i.e. 3T Achieva Philips MR scanner) and exam card used to scan bilingual subjects in Hong Kong were used to scan the monolingual group in order to enhance image comparability.

Data Preprocessing

Bilinguals

High-definition T1 structural scans from the 19 subjects included in the study were pre-processed using SPM8 (www.fil.ion.ucl.ac.uk/spm/). Preprocessing steps included the following: i) first, each image was visually inspected and checked for image-acquisition-related artifacts (e.g. distortions of the magnetic field and movement artifacts) and the origin was reset in order to match the AC-PC (Anterior Commissure – Posterior Commissure) line; ii) Image origin and orientation were aligned to the default Grey Matter (GM) Tissue Probability Map (TPM) using an automated Matlab (The MathWorks, Inc., Natick, Massachusetts, United States) script; iii) Reoriented images were segmented into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) maps according to the VBM8 tissue probability maps and registered to the East Asian brains ICBM (International Consortium for Brain Mapping) space template through affine regularization (Mazziotta et al., 2001); iv) GM segmented maps were input into high-dimensional DARTEL to create non linear, modulated, normalized GM images. The segmentation procedure was further refined by a denoising procedure applying a spatially adaptive nonlocal means filter (Manjón et al., 2010) and through a classical Markov random field approach; v) Finally, non linear, modulated, normalized images were smoothed using a Gaussian kernel of 8mm FWHM (full width at half maximum).

Monolinguals

For the 19 monolingual participants that were included in the study, high-definition T1 structural scans were preprocessed using the same procedures carried out on bilingual subjects' images with the only exception that GM maps were segmented on and registered to the European brain ICBM space template.

VBM (Voxel Based Morphometry) statistics

Bilinguals versus Monolinguals

Given the behavioral effects observed for 1) Conflict Effect (CE) relative to the τ component (CE- τ), calculated as the difference between the τ -incong and τ -cong trials, showing a negative correlation with age; and 2) the positive correlation between μ -incong and τ -cong in both groups, we sought for possible differences between the groups in terms of the association between whole-brain GM volume and each of these observed behavioral effects in brain areas showing a significant aging effect in both groups (i.e. brain areas in which grey matter decreases with age).

Smoothed, normalized and modulated GM maps were thus entered in three separate two sample *t*-test models each including two covariates accounting for the interaction “group x behavioral effect” and one covariate coding age for both groups.

First, regions showing an aging effect (i.e. decrease of GM with an increase of age) in the whole sample (i.e. bilingual and monolingual group) were identified through a contrast testing for the negative (i.e. inverse significant correlation) component of the age regressor. Second, the parameter estimates relative to the main effects of the two covariates resulting from the group x effect interaction were plotted in any areas showing an aging effect in order to assess the direction of the relationship between GM and μ -incong, τ -cong and CE- τ for each group and any potential difference between the two groups for each behavioral effect of interest. Results were thresholded at a more liberal $p < .005$ uncorrected at the voxel level ($k = 10$ voxels) in order to highlight any subtle neural counterpart (i.e. GM volume) of the observed behavioral effects on brain areas found to be sensitive to aging and at $p < .05$ FWE clusterwise corrected for NSS (i.e. smoothness non-stationarity) for GM differences between the two groups. Finally, a contrast was also specified in order to test for whole brain GM differences between the two distinct groups independently of behavioral parameters at the same threshold.

Structural Neuroimaging Results

A significant aging effect was found in the right DLPFC (coordinates: $x=34.5$; $y=34.5$; $z=16.5$) for the whole sample (bilinguals and monolinguals). Namely, a negative correlation was found between age and grey matter volume indicating that an increase in age correlates with a decrease in grey matter volume in right DLPFC. Plots of the parameter estimates for the effects of the two covariates for the μ -incong and τ -cong showed a significant inverse correlation between GM densities in the DLPFC. In other words, higher μ values in response to incongruent trials were paralleled with higher τ values in response to congruent trials and were both associated with lower GM in this area only for the monolingual group. Likewise, the conflict effect of the τ component (CE- τ) was also correlated to GM volume in the right DLPFC exclusively for the monolinguals, showing that an increase of the conflict effect for the τ component was associated with an increase in age specifically in this area. Finally, the contrast assessing overall differences in terms of GM between monolingual and bilingual subjects, revealed a significant difference in the anterior cingulate cortex. Namely, bilingual subjects showed higher GMV in an extensive cluster ($k = 6544$ voxels) peaking at $x = 5$, $y = 38$, $z = -8$, but extending to both rostral and dorsal portions of the ACC when directly compared to monolingual peers (see Figure 4).

(INSERT FIGURE 3 OVER HERE)

Discussion

The aim of our study was to investigate whether bilingualism effectively provides a neural and cognitive reserve for the aging population. For this purpose, we investigated the performance of aging bilinguals and monolinguals on the Flanker Task providing a measurement of their attentional control abilities. For both congruent and incongruent trials bilinguals outperformed monolinguals. We then performed an ex-Gaussian analysis on the resulting RTs, which provides the advantage of not losing critical information about the real trend of RTs by removing the “outliers” that may represent the more informative data for cognitive performance. This is particularly important when investigating a population that may potentially deviate from the normal trend, such as elderly subjects. This method has been successfully applied to investigate potential advantages of young bilinguals in executive control tasks such as the Flanker task (Calabria et al., 2011). Finally, we correlated our behavioral findings to GM volumes of the brains of our aging subjects in order to investigate if cognitive performance predicts GM volumes specifically in brain areas affected by aging. We first discuss the behavioral findings and then discuss the structural neuroimaging findings and finally address the implications of the present findings for healthy aging.

The bilingual advantage as investigated by the ex-Gaussian analysis

The τ component is thought to reflect more controlled processing which under present task conditions involves inhibitory control (i.e. inhibiting the irrelevant stimulus). In contrast, the μ component has been associated with more automatic processes e.g. translating the decision into motor output (Balota & Spieler, 1999), and hence, the speed of response. According to this hypothesis it is plausible that between-group differences in cognitive efficiency may appear mainly in the more demanding part of the process, i.e., the Tau component, if one is to believe that the bilingual advantage resides in more efficient inhibitory control (Abutalebi & Green, 2007). However, Calabria et al. (2011) in a large behavioral study report that bilinguals have an

advantage on both components μ and τ , for incongruent and congruent trials and also for the conflict effect.

Here we report faster values on incongruent trials for the τ component for bilinguals. As aforementioned, this suggests that the bilingual advantage manifests itself when inhibitory control is required. In order to make the correct response, participants have to inhibit interference from distractors and focus on the correct response. These results support the notion of more efficient attentional control in bilinguals (Abutalebi et al., 2012). However, bilinguals also showed significantly faster responses on congruent trials for the μ component and a trend towards significance for incongruent trials on the μ component, reflecting faster processing on more automatic aspects of performance. These findings suggest that the bilingual advantage may not only be limited to controlled processing.

Whilst largely supporting the findings from Calabria et al. (2011), we did not observe significant differences in the conflict effect between the two groups in μ nor in τ . This may be due to the smaller sample size employed in the present study. Alternatively, it is possible that this behavioral effect is not present in an aging population as compared to the young adults tested in the Calabria et al. (2011) investigation. However, we tend to rule out the latter possibility since there is good evidence showing that the magnitude of the bilingual advantage on the conflict effect is particularly evident for elderly bilinguals as compared to younger subjects (see Bialystok et al., 2005). As a further possibility, Calabria et al. (2011) analyzed data from a condition in which there were 75% congruent trials and only in this condition was there a difference in CE between monolinguals and bilinguals. Using 75% congruent trials makes the task harder and may put more pressure on the attentional control system, which is more likely to reveal differences between groups. In the present study, we employed 33% of each type of trial (congruent, incongruent and neutral) and this may be the reason no such effect was observed here.

Similar to the Calabria et al (2011) study, we also correlated the magnitudes of μ and τ . These authors reported that for monolinguals μ and τ were positively correlated whereas in the bilingual group they were not, suggesting that these parameters are independent in bilinguals and dependent in monolinguals. The present results, instead, indicate that μ and τ are dependent parameters for both groups of subjects: we, indeed, report a positive correlation between μ -incong and τ -cong for both groups. Positive correlations between μ -incong and τ -incong or μ -cong and τ -cong were found in none of the groups.

How may we interpret the finding of a positive correlation between μ -incong and τ -cong observed for both groups? This finding may be specifically related to the aging population where we suppose that speed of controlled processing resources and speed of motor responses are physiologically diminished when compared to younger subjects and this effect may be particularly evident on the most difficult trials such as μ -incong. Now, increases of μ -incong correlated to increases of τ -cong and one possibility is that while the components may be independent at younger ages due to a better functional segregation of different components of cognitive control, at later stages of life processes of neural compensation may force a convergence between components in the direction of common neural substrates spared by aging effects. This may translate into a behavioral dependency between the measures, which captures a more automatic component of cognitive control (μ -incong) when enacting the response to the most difficult type of trial (incongruent) and the more controlled component of the cognitive control process (τ -cong), when directing response towards the easiest trial (congruent). This translates into a speed-accuracy trade-off effect meaning that in order to respond fast and accurately to an incongruent trial one needs to inhibit the stimulus-response binding more strongly for the easier response (mirroring what happens for a more dominant L1 with respect to a weaker L2 during bilingual language production). This effect may be driven by the dependency

established in aging between the “controlled” component in terms of increased cognitive effort necessary to release inhibition from the congruent response, as reflected by a longer τ component, and the “automatic” component in terms of increased monitoring difficulties to detect conflict (μ) generated on incongruent trials, as reflected by a longer μ component. A dual-component process seems therefore to be unavoidable in aging in order to be cognitively efficient and optimize task performance as a function of response speed and accuracy, which will necessarily push towards more control (τ) to compensate for lower levels of automaticity (μ) according to the different amount of cognitive conflict arising for each trial type.

In light of our findings, we propose that μ and τ are not independent processes, at least for an aging population, but rather dependent processes that are engaged during attentional control during the Flanker task.

Notably, our results also showed significant positive correlations between age and the μ -cong and μ -incong components and significant negative correlation with the conflict effect in the τ component (CE- τ) for monolinguals only. We show that aging specifically affects the performance of monolinguals, i.e., increasing age corresponds to greater values (i.e., slower processing) of μ -cong and μ -incong, and to smaller values of the CE- τ (i.e., less efficient processing). Consistent with this suggestion, several studies show that bilingual older adults perform significantly better than monolingual older adults on a variety of executive function tasks (Bialystok & Craik, 2010b; Bialystok et al., 2004; 2007; Fernandes et al., 2007).

In sum, we agree with Calabria et al. (2011) that the ex-Gaussian approach provides a powerful tool to study the fine-grained cognitive processes underlying the bilingual advantage. Here we provided evidence that in aging this advantage is not limited to those conditions where controlled processes such as inhibition are at play (i.e., τ component) but also for those

conditions where automatic responses such as overall speed of processing are needed (i.e., μ component). Strikingly, the latter process and the conflict effect are affected by age only for monolinguals. In the next section, we will discuss the neural correlates of these age-induced effects.

(INSERT FIGURE 4 OVER HERE)

Neural reserve induced by bilingualism

As to the correlation analysis of our structural neuroimaging study, we found a significant aging effect in the right DLPFC for the whole sample of subjects. Indeed, a negative correlation was found between age and grey matter volumes indicating that an increase in age correlates with a decrease in grey matter volumes in this region for all subjects (monolinguals and bilinguals). However, only in monolinguals did the parameter estimates, μ -incong and τ -cong, show a significant inverse correlation for GM densities in the right DLPFC (see Figure 3). Likewise, the conflict effect of the τ component (CE- τ) was correlated to GM volumes in the right DLPFC exclusively for the monolinguals, in the sense that an increase of the conflict effect for the τ component was associated with an increase in age. In other words, this relationship is mirrored in neural terms in an area showing an aging effect (i.e. decrease of GM volumes related to age) for both groups, however a greater positive difference between reaction times to incongruent and congruent trials (i.e. more control) is positively correlated with the volume of GM in DLPFC only for the monolinguals but not for the bilingual group. This could be interpreted as showing that monolinguals use more control in order to resolve conflict during aging, while bilinguals probably rely on more automatic processes. Thus, aging induces a decline in performance on conflict resolution that is more prominent in monolinguals with respect to bilinguals and this decline is reflected in decreased gray matter in the DLPFC.

As aforementioned, in the behavioral results we reported that both the μ -incongr and τ -congr correlation and the conflict effect of the τ component (CE- τ) correlated with age but only for the monolinguals. At the brain level, we report that these effects have their neural counterparts specifically in the right DLPFC, an area tightly linked to cognitive control (Miller and Cohen, 2001). Interestingly, the right lateral prefrontal cortex has been previously related to domain-general inhibitory control (Jahfari et al., 2011; Forstmann et al. 2008; Aron, 2011). Following the model of Aron (2011) the DLPFC along with the striatum are involved in selective inhibition. However, recently it was also suggested that the right DLPFC is most active when focusing on speed over selectivity during response inhibition (Smittenaar et al., 2013). Hence, the fact that we found decreased grey matter volumes in this region as being correlated to an increased conflict effect in monolinguals only, suggests that during aging bilinguals have a greater neural reserve specifically in this area. Of note, the DLPFC is usually among the first areas in the brain to be vulnerable to the effects of age as reported in several structural neuroimaging studies (Fjell et al., 2009, Kalpouzos et al., 2009). It follows that the bilingual brain might be better protected against the effects of aging.

Few studies have compared aging bilinguals to matched monolinguals and only one correlated behavioral data to brain structure (Abutalebi et al., 2014). The first study used diffusion tensor imaging (DTI) resting-state analysis to investigate white matter connectivity in 13 late bilingual and 13 monolingual senior adults (mean age=70.5), matched for neuropsychological testing (Luk et al., 2011). Interestingly, bilinguals reported higher values of fractional anisotropy as compared to monolinguals in the corpus callosum, extending anteriorly to the white matter of the frontal lobes. As aforementioned, in the study of Abutalebi et al. (2014) the authors reported in a whole-brain VBM investigation that aging bilinguals have overall increased grey matter compared to matched monolinguals (for age, education, socio-economic status, and cognitive testing) in several brain areas such as bilaterally the temporal poles and the orbitofrontal cortex. Strikingly, for the temporal poles, areas associated to lexical

retrieval and semantics, Abutalebi et al. (2014) also reported that specifically L2 proficiency correlated with grey matter volume in the sense that increasing proficiency predicted increasing grey matter volume.

In the present study, we also report a whole brain analysis between the two groups of subjects (see Figure 4). Bilinguals as compared to monolinguals have increased grey matter in an extensive cluster along the ACC, starting from its dorsal region and extending towards its ventral region. This grey matter finding fits well to the white matter finding of the Luk et al. (2011) investigation that reported increased fractional anisotropy for the corpus callosum and its connections for aging bilinguals. The ACC surrounds the corpus callosum and many of the cingulate projections pass through the corpus callosum (Rash & Richards, 2001). It is therefore plausible that the ACC is a good candidate for the grey matter counterpart of Luk et al.'s (2011) white matter findings. The ACC is an important component in the neural circuit mediating cognitive control and one intimately tied to monitoring conflicting information (Carter et al., 1999; Botvinick et al., 2004). Abutalebi and Green (2007) have also postulated that the ACC is a common locus for language control and resolving non-verbal cognitive control and that bilinguals would use this brain area more often than monolinguals because of the constant need to control two languages. Indeed, it has recently been reported that the ACC is tuned specifically in bilinguals (Abutalebi et al., 2012). These authors reported an experience-dependent effect in the ACC: young bilinguals use this structure more efficiently than monolinguals to monitor non-linguistic cognitive conflicts, and importantly behavioral measures such as the conflict resolution in the Flanker task also correlated positively with local grey matter volume. The present finding of increased grey matter volumes in the ACC in aging bilinguals may provide further evidence of the neural benefits induced by the lifelong use of two languages.

Cognitive reserve classically describes the disconnect between brain status and behavioral performance in dementia i.e. individuals with a high degree of cognitive reserve can endure

greater loss of brain tissue before showing clinical symptoms when compared to individuals with less cognitive reserve (Stern et al., 1994). Cognitive reserve can thus be distinguished from neural reserve. The evidence here shows that bilinguals are indeed faster on the Flanker task than monolinguals and have a significant advantage for the μ component on congruent trials and for the τ component on incongruent trials. Although we have related the effects to GMV differences in the ACC, we have no direct evidence of an association between the behavioural and neural effects. Indeed, it is quite possible that the bilingual RT advantage (showing cognitive reserve) is independent of the differences in brain volume (showing neural reserve) in our healthy sample and in bilingual patients with dementia. Functional studies are required to confirm an association between behavioural performance on the Flanker task and neural activity in the ACC for both bilinguals and monolinguals before we can be certain that cognitive reserve results from greater GMV volume. Bilingualism could enhance cognitive reserve through the development of more compensatory strategies over the lifespan e.g. use of multiple linguistic processes such as reading and writing. Although we favor a neural explanation of the present results, the ability to optimize or maximize performance through differential recruitment of alternative strategies might reflect cognitive efficiency rather than neural reserve per se.

In conclusion, we have shown that during aging bilingualism may provide a cognitive reserve as testified by our behavioral findings assessed with the ex-Gaussian analysis. Aging affected the performance of bilinguals less on the Flanker task. Likewise, at the neural level, aging-induced effects on performance correlated with decreased grey matter density but only for monolinguals; and, overall bilingualism was associated with increased grey matter along the ACC. Taken together, these structural changes may protect against cognitive decline associated with aging and supports the notion that bilingualism acts as a neural reserve, protecting against the onset of cognitive decline.

References

- Abutalebi, J., Brambati, S. M., Annoni, J. M., Moro, A., Cappa, S. F., & Perani, D. (2007). The neural cost of the auditory perception of language switches: An event-related functional magnetic resonance imaging study in bilinguals. *The Journal of Neuroscience*, *27*(50), 13762-13769.
- 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*(9), 2076-2086.
- Abutalebi, J., Rosa, P. A. D., Castro Gonzaga, A. K., Keim, R., Costa, A., & Perani, D. (2013a). The role of the left putamen in multilingual language production. *Brain and Language*, *125*(3), 307-315.
- Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013b). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*, *49*(3), 905-911.
- 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.
- Aron, A.R. (2011). From reactive to proactive and selective control: developing a richer model for stopping inappropriate responses. *Biol Psychiatry*, *69*, 55–68.
- 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*(22), 1938-1944.
- Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: beyond measures of central tendency. *Journal of Experimental Psychology: General*, *128*(1), 32-55.
- Balota, D. A., Tse, C. S., Hutchison, K. A., Spieler, D. H., Duchek, J. M., & Morris, J. C. (2010). Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in Stroop Color Naming. *Psychology and Aging*, *25*(1), 208-218.
- Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychology and Aging*, *19*(2), 290-303.
- Bialystok, E., Craik, F. I., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, *24*(1), 40-49.
- Bialystok, E., Craik, F. I., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1341-1354.

- Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia*, 45(2), 459-464.
- Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10(3), 89-129.
- Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010a). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*, 11(4), 485-508.
- Bialystok, E., & Craik, F. I. (2010b). Cognitive and linguistic processing in the bilingual mind. *Current directions in psychological science*, 19(1), 19-23.
- Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in cognitive sciences*, 16(4), 240-250.
- Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8(12), 539-546.
- Calabria, M., Hernández, M., Martin, C. D., & Costa, A. (2011). When the tail counts: the advantage of bilingualism through the ex-Gaussian distribution analysis. *Frontiers in Psychology*, 2: 250.
- Carter, C. S., Botvinick, M. M., & Cohen, J. D. (1999). The contribution of the anterior cingulate cortex to executive processes in cognition. *Reviews in the Neurosciences*, 10(1), 49-58.
- Costa, A., Hernandez, M., & Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*. 106, 59-86.
- Craik, F. I., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease Bilingualism as a form of cognitive reserve. *Neurology*, 75(19), 1726-1729.
- Della Rosa, P. A., Videsott, G., Borsa, V. M., Canini, M., Weekes, B. S., Franceschini, R., & Abutalebi, J. (2013). A neural interactive location for multilingual talent. *Cortex*, 49(2), 605-608.
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of cognitive neuroscience*, 14(3), 340-347.
- Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage*, 26(2), 471-479.
- Fernandes, M. A., Craik, F., Bialystok, E., & Kreuger, S. (2007). Effects of bilingualism, aging, and semantic relatedness on memory under divided attention. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 61(2), 128.
- Fjell, A.M., Westlye, L.T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., et al., (2009). High Consistency of Regional Cortical Thinning in Aging across Multiple Samples. *Cerebral Cortex*, 19, 2001–2012.

- Forstmann, B.U., Dutilh, G., Brown, S., Neumann, J., Von Cramon, D.Y., Ridderinkhof, K.R., & Wagenmakers, E.J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences*, 105, 17538-17542.
- Garibotto, V., Borroni, B., Kalbe, E., Herholz, K., Salmon, E., Holtorf, V., ... & Perani, D. (2008). Education and occupation as proxies for reserve in aMCI converters and AD FDG-PET evidence. *Neurology*, 71(17), 1342-1349.
- Gunning-Dixon, F. M., Brickman, A. M., Cheng, J. C., & Alexopoulos, G. S. (2009). Aging of cerebral white matter: a review of MRI findings. *International journal of geriatric psychiatry*, 24(2), 109-117.
- 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(4), 410-423.
- Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*, 109(2), 340.
- Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Behavior Research Methods, Instruments, & Computers*, 28(3), 427-445.
- Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review*, 9(2), 394-401.
- Heathcote, A., Brown, S., & Cousineau, D. (2004). QMPE: Estimating Lognormal, Wald, and Weibull RT distributions with a parameter-dependent lower bound. *Behavior Research Methods, Instruments, & Computers*, 36(2), 277-290.
- Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, 18(4), 625-658.
- Jackson, J. D., Balota, D. A., Duchek, J. M., & Head, D. (2012). White matter integrity and reaction time intraindividual variability in healthy aging and early-stage Alzheimer disease. *Neuropsychologia*, 50(3), 357-366.
- Jahfari, S., Waldorp, L., van den Wildenberg, W.P., Scholte, H.S., Ridderinkhof, K.R., & Forstmann, B.U. (2011). Effective connectivity reveals important roles for both the hyperdirect (fronto-subthalamic) and the indirect (fronto-striatal-pallidal) fronto-basal ganglia pathways during response inhibition. *Journal of Neuroscience*, 31: 6891-6899.
- Jernigan, T. L., Archibald, S. L., Fennema-Notestine, C., Gamst, A. C., Stout, J. C., Bonner, J., & Hesselink, J. R. (2001). Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiology of aging*, 22(4), 581-594.

- Kalpouzos, G., Chetelat, G., Baron, J.C., Landeau, B., Mevel, K, et al., (2009). Voxel-based mapping of brain gray matter volume and glucose metabolism profiles in normal aging. *Neurobiology of Aging*, 30 , 112–124.
- Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, 106(16), 6556-6560.
- Luce, R. D. (1986). Response times. *New York: Oxford University Press*.
- Luk, G., Bialystok, E., Craik, F. I., & Grady, C. L. (2011). Lifelong bilingualism maintains white matter integrity in older adults. *The Journal of Neuroscience*, 31(46), 16808-16813.
- Madden, D. J., Bennett, I. J., & Song, A. W. (2009). Cerebral white matter integrity and cognitive aging: contributions from diffusion tensor imaging. *Neuropsychology Review*, 19(4), 415-435.
- Manjón, J. V., Coupé, P., Martí Bonmatí, L., Collins, D. L., & Robles, M. (2010). Adaptive non local means denoising of MR images with spatially varying noise levels. *Journal of Magnetic Resonance Imaging*, 31(1), 192-203.
- Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, 16(5), 798-817.
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., ...& Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1412), 1293-1322.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Neurolinguistics: structural plasticity in the bilingual brain. *Nature*, 431(7010), 757-757.
- Miller, E., & Cohen, J. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232-258.
- Pfefferbaum, A., Adalsteinsson, E., & Sullivan, E. V. (2005). Frontal circuitry degradation marks healthy adult aging: evidence from diffusion tensor imaging. *Neuroimage*, 26(3), 891-899.
- Posner, M. I. & Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci*, 13, 25-42.
- Rash, B.G., & Richards, L.J. (2001). A role for cingulate pioneering axons in the development of the corpus callosum. *J Comp Neurol.*, 434(2), 147-57.
- 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(11), 1676-1689.

- Scarmeas, N., Stern, Y., Tang, M. X., Mayeux, R., & Luchsinger, J. A. (2006). Mediterranean diet and risk for Alzheimer's disease. *Annals of neurology*, 59(6), 912-921.
- Schweizer, T. A., Ware, J., Fischer, C. E., Craik, F. I., & Bialystok, E. (2012). Bilingualism as a contributor to cognitive reserve: evidence from brain atrophy in Alzheimer's disease. *Cortex*, 48(8), 991-996.
- Smittenaar, P., Guitart-Masip, M., Lutti, A., & Dolan, R.J. (2013). Preparing for Selective Inhibition within Frontostriatal Loops. *Journal of Neuroscience*, 33(46): 18087–18097.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, 6(2), 174-215.
- Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance*, 22(2), 461-479.
- Stern, Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer Disease & Associated Disorders*, 20(2), 112-117.
- Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, 47(10), 2015-2028.
- Tse, C. S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, 24(3), 300.
- Valenzuela, M. J., & Sachdev, P. (2006). Brain reserve and dementia: a systematic review. *Psychological medicine*, 36(04), 441-454.
- Valian, V. Bilingualism and Cognition. (2015) *Bilingualism: Language and Cognition* (in press)
- Verhaeghen, P., & Hoyer, W. J. (2007). Aging, focus switching, and task switching in a continuous calculation task: evidence toward a new working memory control process. *Aging, Neuropsychology, and Cognition*, 14(1), 22-39.
- 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.

Tables

Table 1. Socio-demographical data description of the experimental samples (*i.e.* bilinguals and monolinguals). Mean, Standard Deviation (st dev) and range values are provided for each variable (*i.e.* Age, Education, Mini Mental State Examination and Socio Economic Status) along with group statistical comparability indexes (t tests, *p* values).

<i>Descriptives</i>		<i>Bilinguals</i> (N=19 8 M / 11 F)	<i>Monolinguals</i> (N=19 9 M / 10 F)	T Test p Value
<i>Age</i>	<i>mean</i>	61.68	60.93	.680
	<i>st dev</i>	5.34	5.81	
	<i>range</i>	55 : 75	49 : 75	
<i>Education</i>	<i>mean</i>	13.76	13.16	.626
	<i>st dev</i>	3.98	4.86	
	<i>range</i>	8 : 24	5 : 25	
<i>MMSE</i>	<i>mean</i>	28.63	28.95	.416
	<i>st dev</i>	0.9	1.03	
	<i>range</i>	27 : 30	27 : 30	
<i>SES</i>	<i>mean</i>	22.18	22.55	.813
	<i>st dev</i>	5.39	4.12	
	<i>range</i>	14 : 36	15 : 37	

Table 2. Language competence assessment and Flanker task descriptive statistics (mean and standard deviation) for bilingual and monolingual subjects are reported. Accuracy scores (% Accuracy) are reported for translation task (Translation), naming task performed with the mother tongue (L1 Naming) and with the second language (L2 Naming) and for Flanker task. Language competence assessment scores are also provided for hours of daily exposure (hpd) to the second language (L2 exposure) and age of second language acquisition (AoA L2, expressed in years). (*See “Material and Methods” section for a comprehensive description of the tests.*)

<i>Descriptives</i>		<i>Bilinguals</i> (N=19 8 M / 11 F)	<i>Monolinguals</i> (N=19 9 M / 10 F)
<i>L1 Naming</i> (% Accuracy)	<i>mean</i>	82	99
	<i>st dev</i>	1.0	0.01
<i>L2 Naming</i> (% Accuracy)	<i>mean</i>	66	-
	<i>st dev</i>	1.0	-
<i>Translation</i> (% Accuracy)	<i>mean</i>	52.47	
	<i>st dev</i>	10.07	
<i>L2 Exposure</i> (hpd)	<i>mean</i>	2.92	-
	<i>st dev</i>	3.40	-
<i>AoA L2</i> (yy)	<i>mean</i>	12.68	-
	<i>st dev</i>	10.71	-
<i>Flanker</i> (% Accuracy)	<i>mean</i>	99.21	99.23
	<i>st dev</i>	0.01	0.01

Table 3. Mean, SD and 95% confidence interval values relative to the three accuracy performance criteria chosen for inclusion in the ex-gaussian analysis for the final sample of 38 subjects (i.e. 19 bilinguals and 19 monolinguals).

<i>Criteria</i>	<i>Group</i>	<i>Mean</i>	<i>SD</i>	<i>95% Confidence Interval (expressed in %)</i>
<i>1. Global accuracy</i>	<i>BIL (n=19)</i>	99,21%	1,05%	97.15 - 101.26
	<i>MONO (n=19)</i>	99,23%	1,09%	97.10 - 101.37
	<i>Total (n=38)</i>	99,22%	1,05%	97.15 - 101.28
<i>2. Accuracy difference (Congruent - Incongruent)</i>	<i>BIL (n=19)</i>	0,90%	1,75%	-2.53 - 4.34
	<i>MONO (n=19)</i>	0,82%	1,20%	-1.54 - 3.18
	<i>Total (n=38)</i>	0,86%	1,48%	-2.05 - 3.77
<i>3. Accuracy Difference (sessions) (Congruent - Incongruent)</i>	<i>BIL (n=19)</i>	0,51%	3,69%	-6.73 - 7.75
	<i>MONO (n=19)</i>	0,00%	2,55%	-5.00 - 5.00
	<i>Total (n=38)</i>	0,26%	3,14%	-5.90 - 6.41

* Mean accuracy percentage scores for the first criteria were very high, thus upper CI boundary values exceed 100%

Figure legends

Figure 1. Stimuli and procedures of the Flanker task. See the *Experimental Procedures* section for further details about the task.

Figure 2. Differences in reaction times (RTs) between bilinguals (solid line) and monolinguals (dotted line) on Congruent and Incongruent trials for the τ (tau, left panel) and μ (Mu, right panel) components of the ex gaussian analysis.

Figure 3. The figure reports the aging effect in the right DLPFC (coordinates: $x=34.5$; $y=34.5$; $z=16.5$) found for both bilinguals and monolinguals showing a negative correlation between age and grey matter volumes (left panel). The right panels report the plots of the parameter estimates for τ congruent (top) and μ incongruent (bottom) showing a significant inverse correlation with GM volumes in the DLPFC only for the monolingual group.

Figure 4. Comparison of grey matter whole brain volumes between bilinguals and monolinguals. Bilingual subjects showed higher values for grey matter volume when compared to their monolingual peers in a cluster peaking at $x=5$, $y=38$, $z=-8$ and extending throughout the Anterior Cingulate Cortex (ACC), bilaterally. The figure shows the cluster extension in the left (first and second sagittal slices) and right (third and fourth sagittal slices) ACC added with a coronal plane displaying slices cutout.