Northumbria Research Link

Northumbria
University
NEWCASTLE

Running Head:

BILINGUAL EXPERIENCE, EXECUTIVE CONTROL AND NEUROPLASTICITY

Individual differences in bilingual experience modulate executive control network and performance: behavioral and structural neuroimaging evidence*

Federico Gallo[1,2], Nikolay Novitskiy[3,4], Andriy Myachykov[1,5], & Yury Shtyrov[1,6]

[1]Centre for Cognition and Decision Making, Institute for Cognitive Neuroscience, National Research University Higher School of Economics, Moscow, Russian Federation
[2]Centre for Neurolinguistics and Psycholinguistics, Vita-Salute San Raffaele University, Milan, Italy
[3]Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong, China
[4]Brain and Mind Institute, The Chinese University of Hong Kong, Hong Kong, China
[5]Northumbria University, Newcastle-upon-Tyne, United Kingdom
[6]Center of Functionally Integrative Neuroscience, Aarhus University, Aarhus, Denmark

*Aknowledgments:

The present study was supported by Russian Science Foundation Grant (project №19-18-00550) to the National Research University Higher School of Economics.

Address for correspondence:

Federico Gallo

Centre for Cognition and Decision Making, Institute for Cognitive Neuroscience,

National Research University Higher School of Economics

Krivokolenniy Pereulok, 3, Entrance 2, Moscow, Russian Federation

+7 985 747 39 76

Email: fgallo@hse.ru

Keywords: bilingualism, bilingual experience factors, executive control, structural MRI, region-based-morphometry

Abstract

Dual/multiple language use has been shown to affect cognition and its neural substrates, although the replicability of such findings varies, partially due to neglecting the role of interindividual variability in bilingual experience. To address this, we operationalized the main bilingual experience factors as continuous variables, investigating their effects on executive control performance and neural substrate deploying a Flanker task and structural magnetic resonance imaging. First, higher L2 proficiency predicted better executive performance. Second, neuroimaging results indicated that bilingualism-related neuroplasticity may peak at a certain stage of bilingual experience and eventually revert, possibly following functional specialization. Importantly, experienced bilinguals optimized behavioral performance independently of volumetric variations, suggesting a degree of performance gain even with lower GMV. Hence, the effects of bilingualism on cognition may evolve with experience, with improvements in functional efficiency eventually replacing structural changes. We conclude that individual differences in bilingual experience modulate cognitive and neural consequences of bilingualism.

1. Introduction

Existing evidence suggests that, in a bilingual brain, two language representations are simultaneously active and compete for selection during language use (e.g., Kaushanskaya & Marian, 2007; Kroll, Bobb & Hoshino, 2014; Novitskiy, Myachykov & Shtyrov, 2019). Thus, bilinguals need to exercise constant control over their two languages in order to constrain competition and successfully carry out communication. In other words, cognitive mechanisms of selection, inhibition and switching must be continuously engaged in order to manage potential interferences between the two language codes. To accomplish such a task, bilinguals rely on a set of abilities referred to as *language control* (Abutalebi & Green, 2007; Green & Abutalebi, 2013). Bilingual language control relies on a neural network comprising cingulo-fronto-parietal cortical areas as well as subcortical structures, largely overlapping with the domain-general executive control[1] network (Abutalebi & Green, 2016). It is argued that, due to the overlap between the mechanisms and brain areas implied in bilingual language control and general executive control, the continuous training experienced by bilingual users may eventually impact their domain-general executive functioning, both at the cognitive and at the neural level (e.g., Bialystok, 2017; Kroll, Dussias, Bice & Perotti 2015). Indeed, several studies reported volumetric changes in areas associated with bilingual language control, with highly proficient bilinguals showing greater grey matter density than age-matched monolinguals (for a review see Li, Legault & Litcofsky, 2014). Similar bilingualism-induced neuroplastic changes have been reported for the white matter tracts connecting key areas of the language control network, such as the inferior and superior longitudinal fasciculi and the inferior fronto-occipital fasciculus (e.g., Gold, Johnson & Powell, 2013; Luk, Bialystok, Craik & Grady, 2011a). These and similar studies have been taken to suggest specific structural brain changes in the executive network as a result of multiple language use.

In similar vein, bilingual experience has also been argued to affect executive control at the behavioral level, with several studies reporting highly proficient bilinguals outperforming monolinguals on a number of executive tasks, across different age groups (for a review, see Bialystok, Craik & Luk, 2012). Nevertheless, considerable variability between the results can be observed in bilingual research. Importantly, several studies failed to replicate findings of enhanced executive performance in bilinguals (e.g., Paap, Johnson & Sawi, 2015) leading some to question the existence of any beneficial effect of bilingualism on cognition (e.g., Lehtonen, Soveri, Laine, Järvenpää, de Bruin & Antfolk, 2018). Similarly, contradicting evidence accompanies the research on the consequences of bilingualism for brain structure, particularly in terms of the effect localization (see García-Pentón, Fernandez Garcia, Costello, Duñabeitia, & Carreiras, 2016). Such conflicting evidence has been mainly attributed to inconsistencies in how bilingualism is defined (e.g., Luk & Bialystok, 2013; Mishra, 2015; Surrain & Luk, 2019). Indeed, although interindividual variability is intrinsic to such a multifaceted phenomenon as bilingualism, the tendency still persisting in the literature is to reduce diverse linguistic profiles to a dichotomous categorization (bilinguals vs. monolinguals). This approach often neglects important differences within heterogeneous groups and downplays the role of individual differences in profiles of bilinguals in modulating neuroplastic and cognitive changes (e.g., Luk and Bialystok, 2013; Bialystok, 2016). In other words, neglecting the fact that bilingualism is a continuous rather than a binary variable may make study designs insensitive to its influence on other neurocognitive functions.

At the same time, existing research indicates that individual differences in bilingual experience factors (BEFs) may play an important role in modulating the effects of dual language use on cognition and neuroplasticity. Regarding the impact of individual BEF differences on executive control at the behavioral level, previous studies mainly focused on the role of second language (L2) proficiency. When comparing bilinguals with differing levels of L2 proficiency, higher L2 proficiency has generally been found to be associated with a greater improvement in executive control. This result has been reported as better conflict resolution performance in a lateralized attention network task (Tao, Marzecová, Taft, Asanowicz & Wodniecka, 2011) and in both canonical (Iluz-Cohen & Armon-Lotem, 2013) and oculomotor (Singh & Mishra, 2012; 2013) versions of the Stroop task, in reduced conflict effect in the Flanker task (Luk, De Sa & Bialystok, 2011b; Sorge, Toplak & Bialystok, 2017; Novitskiy, Shtyrov & Myachykov, 2019) and better endogenous attention disengagement in Posner's cueing paradigm (Mishra, Hilchey, Singh & Klein, 2012). The rationale behind relating higher L2 proficiency to better executive control follows from the observation that bilinguals who are highly fluent in L2 experience a stronger cross-linguistic interplay between their two language systems. It has indeed been reported that highly proficient bilinguals automatically activate L2 lexical items when processing L1-specific targets whereas low proficient ones do not (Blumenfeld & Marian, 2007). Higher L2 proficiency has also been linked to stronger parallel activation of lexicons (e.g., Van Hell & Dijkstra, 2002), increasing overlap in the time course of L1 and L2 lexicon activation (Guo & Peng, 2006) as well as unconscious activation of translation equivalents in the non-target language (e.g., Wu & Thierry, 2010).

Some studies also investigated the impact of individual- and group-level BEF differences on neuroplasticity of the language/executive control network areas. In a landmark study, Mechelli et al. (2004) reported that increases in the grey matter volume (GMV) of the left inferior parietal lobule of bilingual users are positively associated with L2 proficiency. Similar results have been replicated in studies deploying various tasks as proxies of L2 proficiency. Grogan, Green, Ali, Crinion and Price (2009), for example, reported a positive association between GMV of the bilateral caudate nucleus and performance on a phonemic fluency task. The same group (Grogan et al., 2012) also showed a positive correlation between left inferior frontal gyrus's GMV and scores in lexical decision and verbal fluency tasks in L2. Similarly, Pliatsikas, Johnstone and Marinis (2014) reported increases in GMV of the cerebellum associated with higher speed of processing of regular verb morphology in L2. Abutalebi et al. (2014) reported an association between increases in GMV of the left temporal pole and performance in an L2 picture naming task. Finally, Hervais-Adelman, Egorova and Golestani (2018) reported neuroplastic changes in a sample of multilingual speakers in the shape and volume of bilateral caudate nuclei predicted by a composite score accounting for L2 age of acquisition (AoA) and proficiency of each of the languages spoken by the participants.

Similar results also emerged concerning L2 exposure and immersion (i.e., the length of residence in an L2-speaking country). For instance, in their abovementioned 2014 study, Pliatsikas and colleagues reported a positive association between L2 exposure and GMV of the putamen, bilaterally. The same group reported intriguing results in a more recent study (Pliatsikas, DeLuca, Moschopoulou & Saddy, 2017) investigating the effect of immersion on grey matter density of subcortical regions of the bilingual language control network. Increased GMV, as compared to monolinguals, was registered in highly-immersed bilinguals in the bilateral putamen, right thalamus and bilateral globus pallidus, whereas higher GMV in the left caudate nucleus were reported for the low-immersion group, suggesting that L2 exposure modulates the impact of bilingualism on the brain.

Taken together, these findings highlight the importance of considering the role of individual BEF differences in fostering neuroplastic and cognitive changes in bilinguals. It is also worth noting that, beside the tendency to prefer group comparisons to evaluation of continuous language background variables, research on the cognitive consequences of bilingualism still tends to focus on early/balanced over late/unbalanced bilinguals. This “sampling bias” disregards the fact that late/unbalanced bilinguals represent a steadily increasing majority of the world’s bilingual population due to globalization and global migration processes, most expressed in the ever-increasing diffusion of English as an L2 in the globalized society. Another methodological issue in bilingual research has been the relative scarcity of studies investigating the relationship between behavioral and neural consequences of bilingualism, resulting in the lack of specificity. As suggested by Del Maschio and colleagues (2018) in one of the few studies adopting such a practice, investigating the relationship between bilingualism-induced neuroplastic changes and behavioral outcomes might inform us on *qualitative,* in addition to *quantitative,* bilingualism-induced changes in cognition (see also section 4 below for further discussion).

Given these premises and in line with the recent trends in bilingual research (e.g., Del Maschio et al., 2019; DeLuca, Rothman, Bialystok & Pliatsikas, 2019; Hervais-Adelman et al., 2018), we designed and conducted a study aimed at filling these gaps. Here, we operationalized the main aspects of bilingual experience (i.e., L2 AoA, exposure and proficiency) as continuous variables in an attempt to further investigate the effects of bilingualism on cognition and neuroplasticity. Hence, our main purpose was to investigate the effects of individual differences in BEFs on the executive performance and its neural substrate in a sample of late unbalanced Russian-English bilinguals. We estimated individual differences in bilingual experience on a range of indicators, obtained a measure of executive performance using a Flanker task, and analyzed the relationship between these measures and GMV of the executive control network, obtained using region-based morphometry. Based on previous findings, we hypothesized that:

i) Higher levels of bilingual experience would correspond to better executive control performance at the behavioral level (e.g., Luk, et al., 2011b; Sorge et al., 2017; Novitskiy et al., 2019);

ii) Continuous (rather than categorical) BEF measurement could illuminate the trajectory of bilingualism-related neuroplasticity. In particular, we expected to register effects in line with Pliatsikas' Dynamic Restructuring Model (Pliatsikas, 2019), suggesting a reverse u-shaped relationship between GMV and bilingual experience in the targeted brain regions (i.e., an increase, peak and subsequent decrease of GMV as a function of an increasing bilingual expertise);

iii) Continuously measured BEFs paired with the combination of structural neuroimaging and behavioral analyses would illuminate the role of bilingual experience in modulating the brain structure-behavior relationship in the executive control network. In particular, we expected higher levels of bilingual experience to mitigate such a relationship due to enhancements in the efficacy and the flexibility of the executive control network fostered by the bilingual experience in the first place (Del Maschio et al., 2018).

2. Materials and methods

2.1 Participants

Twenty-two Russian-English speaking participants (9 males; mean age = 22.95, SD ± 4.38) were recruited mostly from the population of students of the HSE Department of Psychology. All participants acquired English as an L2 formally through instruction at school, although at different ages. All participants were right-handed, as established by the Edinburgh Handedness Inventory scale (Oldfield, 1971). No participant had a history of neurologic or psychiatric illnesses. Socio-demographic variables – age, educational attainment and socio-economic status (SES) – were collected for all participants using the MacArthur Scale of Subjective Social Status (MacArthur Foundation, 2007). The annual household income bands, here used as a proxy of SES, were adapted to Russian Federation standards based on the European Social Survey 2016 (European Social Survey, 2016). Participants' fluid intelligence was also assessed, using a subset of the Raven's Standard Progressive Matrices for adults (Raven, Raven & Court, 2000). Details on demographic and cognitive measures are reported in Table 1. The study was approved by the local research ethics committee, and written informed consent was obtained from all participants.

2.2 Bilingual-experience-factors: L2 age of acquisition, exposure, and proficiency

L2 AoA and exposure were measured using the Russian version of the Language Experience and Proficiency Questionnaire (LEAP-Q), an established tool for assessing language background of bi- and multilingual populations (Marian, Blumenfeld & Kaushanskaya, 2007). The LEAP-Q was implemented in the NBS Presentation® software (Version 18.1, Neurobehavioral Systems, Inc., 2020) such that each question of the questionnaire appeared as one computer screen. The AoA and daily exposure to English measurements were self-rated by participants. To obtain an objective measure of English proficiency, we deployed a computerized custom-design L2-vocabulary task, implemented in the same version of the NBS Presentation® software as detailed above. The vocabulary task, including 146 English words of differing lexical frequencies, consisted in an unspeeded forced choice in which the participant was presented with an L2 word and three L1 alternatives, all in one column. The participant had to select the semantically matching L1 alternative by pressing 1, 2, or 3 on the keyboard. Feedback was provided at the end of each trial for both correct and incorrect responses (for more details of the task, please see Novitskiy, Shtyrov and Myachykov, 2019). While the ability to accurately match translations between L2 and L1 cannot be taken as a *comprehensive* measure of bilingual proficiency, it has been previously used as *a proxy* of the latter (e.g. Perani et al., 2003; Abutalebi et al., 2013; Abutalebi, et al., 2014; Abutalebi et al., 2015). Our choice of this task was motivated by the fact that vocabulary tests are the easiest to be quantified (Treffers-Daller, 2019), that they are relatively quick and easy to develop and score, and that they are quite reliable since they are relatively objective as suggested, for example, by Milton (2009) who also emphasizes that vocabulary tasks (both receptive and productive) correlate well with grades of many examination frameworks, such as IELTS or CEFR. Details of the linguistic measures are reported in Table 1.

<Insert Table 1 about here>

2.3 Behavioral assessment

All participants performed the Flanker task within the standard ANT set-up (Fan, McCandliss, Sommer, Raz & Posner, 2002). Participants were seated in an electrically shielded and acoustically dampened chamber. Experimental stimuli were presented on a 75 cm-diagonal computer screen. Target trials proceeded from the initial fixation point presented at the center of the screen for 400 ms, followed by a row of five horizontal black lines with arrow heads pointing to the left or to the right for 1700 ms. Participants were instructed to detect and signal the direction of the central target arrow by pressing the left or right arrow button on the keyboard as fast as possible. Targets appeared with additional arrows flanked to the same direction as the target arrow (→→→→→) (i.e., congruent condition), with surrounding arrows flanked to the opposite direction of the target (←←→←←) (i.e., incongruent condition), or with surrounding neutral lines (-- -- → -- --) (i.e., neutral condition). Congruent flankers cue the correct response and are typically associated with faster and more accurate task performance; incongruent flankers provide conflicting visual information with regard to the correct response and generally yield performance decline associated with lower accuracy and increasing RT; the neutral condition supports neither the correct nor the incorrect response. Congruent, incongruent, and neutral trials were presented in a pseudo-randomized order during three runs with 96 items (32 for each condition) per run. Prior to the experiment, participants had a practice run consisting of 24 pseudo-randomized trials.

2.4 MRI data acquisition and preprocessing

T1-weighted images were acquired with a Philips Intera 1.5T MRI scanner, using the following parameters: TR = 25 ms, TE = 4.6 ms; flip angle = 30, FOV = 240x240, resolution = 1x1x1 mm, matrix = 256, TA = 5.35 min, mode = 3DFFE, number of slices = 191. The Computational Anatomy Toolbox (CAT12, r1113, Gaser, 2020) within SPM12 (v6906) was used to obtain total amount of GMV within pre-defined brain regions comprising the bilingual language control network (Abutalebi & Green, 2016), by performing region-based morphometry. Images were first visually inspected to check for gross field distortions and movement artifacts; 2 participants had to be discarded for this reason. For each image, the origin was manually set to correspond to the AC-PC (anterior commissure-posterior commissure) line. Then, the following two-steps procedure was used for GMV extraction. In the first step, raw structural images were segmented into Grey Matter (GM), white matter (WM) and cerebrospinal fluid (CSF) images. The segmentation routine implemented in CAT12 utilizes an adaptive Maximum A Posterior (aMAP) technique that reduces the need for *a priori* information about tissue probabilities (see Rajapakse, Giedd & Rapoport, 1997) and also accounts for local variations and inhomogeneities of GM intensity (Dahnke, Ziegler & Gaser, 2012). Following aMAP segmentation, CAT12 also carries out a Partial Volume Estimation (PVE) of mixed tissue-classes (GM-WM and GM-CSF) (Tohka, Zijdenbos & Evans, 2004) that results in a more accurate segmentation by estimating the fraction of pure tissue of each type within each voxel. The segmentation routine was further improved by using a spatial-adaptive non-local means (SANLM) denoising filter in a pre-segmentation step (Manjón, Coupé, Martí-Bonmatí, Collins & Robles, 2010). Following segmentation, the brains of all participants were registered to the ICBM (International Consortium for Brain Mapping) European brain space template by affine regularization. In the final step, GMV values were extracted from the following regions of interest (ROIs) that are constituent parts of the bilingual language control network (see Figure 1) (Abutalebi & Green, 2016): (1) Left and right anterior cingulate cortex (LACC, RACC); (2) left and right caudate nucleus; (3) left and right prefrontal cortex (LPFC, RPFC); (4) left and right inferior parietal lobule (LIPL, RIPL). The extraction was performed using an in-built CAT12 function allowing for the estimation of GMV in non-normalized native space using maximum tissue probability labels derived from the Neuromorphometrics Atlas (2012) (Neuromorphometrics, Inc.). To control for individual differences in brain sizes, Total Intracranial Volume (TIV) was calculated for each participant by summing the native space global volumes of GM, WM and CSF.

<Insert Figure 1 about here>

2.5 Statistical analyses

In order to conduct analyses on the Flanker task's reaction times (RTs), we removed error trials, false start trials with RTs below 100 ms, and outlier trials, i.e., those falling outside two standard deviations from the individual subjects' mean RT values. Neutral condition trials were discarded, as we aimed to investigate the conflict effect, a measure of inhibitory control computed as the RT difference between incongruent and congruent trials. Subsequently, to check for homogeneity in SES, intelligence, and education, we tested for outliers in our sample, again using a two standard deviations threshold. No outliers were detected for these measures. In addition, to control for eventual confounding effects of such factors on executive control performance, we tested a linear mixed effects model including Flanker's RT as the dependent variable, together with main effects of SES, intelligence and education as fixed factors and crossed random effects of trial and participant. No significant effects emerged for any of the predictors.

Effects of individual differences in bilingual experience factors on executive behavioral performance

To examine the effects of the BEF differences on executive behavioral performance, linear mixed effects analyses were conducted using the *lme4* (Bates, Maechler, Bolker & Walker, 2014) and the *lmerTest* (Kuznetsova, Brockhoff & Christensen, 2017) packages in R (R Core Team, 2014). The choice of mixed effects modeling, alongside other advantages, allowed us to increase the number of the data points by assessing executive performance on a trial-by-trial level. Consequently, as the standard Flanker conflict effect, computed as the increase in mean RTs in the incongruent (vs. congruent) condition, could not be deployed, we selected, as a measure of conflict, the contribution of task condition (congruent vs. incongruent) to modulating the effect of our predictors (i.e. the interactions between predictors and task condition). The full model included Flanker's RT as the dependent variable, alongside L2 AoA, L2 exposure, L2 proficiency and task condition, and the interactions between these predictors, as fixed factors, as well as crossed random effects of trial and participant. Subsequently, the best fitting model was obtained by deploying the *step* function implemented in the *lmerTest* package (Kuznetsova et al., 2017), which allows to perform automatized step-down model selection. The best fitting model included main effects of task condition and L2 proficiency as well as the interaction between these two factors.

Effects of individual differences in bilingual experience factors on the language/executive control network

To investigate the role of BEF differences in bilingualism-related neuroplasticity in the language/executive control network, we used the *lm* function implemented in the *stats* package (R Core Team, 2014) to fit linear models with the GMV of each ROI as the dependent variable, and L2 AoA, exposure and proficiency as predictors. TIV was used as a covariate in all the models to control for individual differences in brain size. Subsequently, the best fitting model was obtained by deploying the *step* function implemented in the *stats* package (R Core Team, 2014), again via an automatized backwards stepwise search. For all the significant ROIs (see the Results section below), the best fitting model included main effects of L2 exposure and L2 proficiency, together with their interaction.

Relationships between individual differences in bilingual experience factors and grey matter volume in modulating executive control performance

Lastly, we conducted linear mixed effects analyses to assess whether individual differences in BEFs modulate the relationship between GMV and executive performance. To avoid the risk of overfitting – and the difficulties in the interpretation of results – related to testing a model including 5 continuous variables as predictors, we decided to compute a bilingual index (BI) taking into account the contribution of all three BEFs, i.e., L2 AoA, exposure and proficiency. We used the following logic in order to compute this combined index. We assumed L2 exposure and L2 proficiency to be positively associated with total bilingual experience, and thus contribute positively to the BI, whereas L2 AoA to contribute negatively (the later AoA, the less its contribution). Thus, the BI formula (reported below) included L2 exposure and L2 proficiency in the numerator, and L2 AoA in the denominator. Since our aim was to investigate the contribution of each BEF to modulating the inhibitory control ability, we tested our prediction about directionality of associations by fitting a linear mixed effects model that included Flanker's RT as the dependent variable, main effects of L2 exposure, L2 proficiency, L2 AoA and task condition, and the interactions between task condition and each of the factors as predictors, with trial and participant entered as crossed random effects. Since it is impossible to compute means for congruent and incongruent trials with a trial-by-trial approach, we focused on the interactions between the single BEFs and task condition as a measure of the contribution of individual BEFs to the conflict effect. Our results confirmed our assumptions as the interactions containing L2 exposure and L2 proficiency showed an inverse relationship with RTs (i.e., increasing levels of such factors predicted better executive performance), while the interaction containing L2 AoA revealed the opposite pattern. Once assumptions regarding numerator and denominator positions were confirmed, we proceeded to weigh the contribution of the single BEFs in order to obtain a more precise index. The weights then were obtained from the absolute values of interactions' beta estimates in the abovementioned model. The formula used to compute the BI was thus the following:

$$BI = \frac{(a * \text{L2 exposure}) * (b * \text{L2 proficiency})}{(c * \text{L2 AoA})}$$

where *a*, *b*, and *c* are the absolute values of the interactions' beta estimates: a = L2 exposure*task condition = |-3.029|; *b* = L2 proficiency*task condition = |-8.205|; *c* = L2 AoA*task condition = |3.468|. Of note, previous attempts to build an index accounting for different dimensions of the bilingual experience have been somewhat sparse (e.g., Gollan, Salmon, Montoya & Galasko, 2011), and thus there may be potentially some interest in using our approach to computing a composite BI in other studies. While we deem our strategy to be overall generalizable to different applications, caution should still be exercised with respect to carefully considering the specific experimental hypotheses when deriving the factor weighs. Indeed, an accurate rendering of the contributions of the single BEFs to the total BI cannot disregard study-specific hypotheses and sample characteristics.

Once we computed the BI, we fitted a linear mixed effects model for each ROI, including Flanker's RT as the dependent variable, fixed effects of task condition, BI, GMV, and their interactions as predictors, and crossed random effects of trial and participant.

3. Results

3.1 Effects of individual differences in bilingual experience factors on executive behavioral performance

Mean Flanker RTs and accuracy measures are reported in Table 2. The stepwise model selection procedure returned as the best fitting model that including main effects of task condition and L2 proficiency as well as the interaction between these two factors. A significant task condition by L2 proficiency interaction was registered in the RT data, with higher L2 proficiency scores predicting better executive performance (i.e., lower RTs, differentially in the incongruent condition; F value = -3.926, Pr(>F) = 8.79e-05), as shown in the interaction plot reported in Figure 2.

<Insert Table 2 about here>

<Insert Figure 2 about here>

3.2 Effects of individual differences in bilingual experience factors on the language/executive control network

We registered a significant effect of BEFs on GMV of the bilateral ACC. This effect also exhibited a trend towards significance in the bilateral PFC. For all four ROIs, the stepwise model selection identified, as the best fit, a model including main effects of L2 exposure and L2 proficiency, together with their interaction, as predictors of GMV. A crossed L2 proficiency*L2 exposure interaction significantly predicted GMV in the left ACC (t value = -3.016, Pr(>|t|) = .00869) and right ACC (t value = -3.791, Pr(>|t|) = .001777), approaching significance in the left PFC (t value =-2.087, Pr(>|t|) = .054318) and right PFC (t value = -2.109, Pr(>|t|) = .052204). As shown in the interaction plots reported in Figure 3, higher proficiency was associated with increases in GMV only at lower levels of exposure, and, vice-versa, higher exposure was associated with increases in GMV only at lower levels of proficiency. When both of the predictors reached high values, a decreasing trend in GMV emerged.

<Insert Figure 3 about here>

3.3 Relationships between individual differences in bilingual experience factors and grey matter volume in modulating executive control performance

A significant interaction between individual differences in BEFs and GMV on executive performance emerged for the bilateral caudate and the left PFC. For all these ROI the best fitting model coincided with the full model. In each of the three ROI, a comparable pattern emerged, with BI significantly interacting with GMV and task condition in predicting Flanker RTs. The interaction plot shown in Figure 4 shows that variations in GMV predicted incongruent trials' RTs (i.e., those tapping on inhibitory control) at lower, but not at higher levels of BI. In other words, executive performance appeared to be unrelated to GMV of the left caudate (F value = 4.3636, Pr(>F) = .036787), right caudate (F value = 4.2238, Pr(>F) = .03993), and left PFC (F value = 4.9149, Pr(>F) = .0266913) at higher levels of BI. Overall, higher levels of BI predicted better executive performance (i.e., lower RTs). Lastly, in the face of increasing GMV, performance levels tended to overlap irrespectively of the level of BI.

<Insert Figure 4 about here>

4. Discussion

The challenge of investigating such a complex, multifaceted phenomenon as bilingualism requires going beyond a group-comparison approach. In the present study, we operationalized three main aspects of the bilingual experience – namely AoA, exposure, and proficiency in the L2 – as continuous variables, in the attempt to elucidate the consequences of dual language use for cognition. We tested the contribution of these factors on executive control, both at the behavioral and at the neural levels. Increases in L2 proficiency significantly predicted better executive control performance. Moreover, the interplay between L2 exposure and proficiency was shown to affect bilingualism-induced neuroplastic changes, which increased with increasing L2 fluency and use, albeit plateauing and eventually regressing at further increasing levels of bilingual competence and exposure. Furthermore, our results indicated that a bilingual experience index, taking into account all three BEFs, modulated the relationship between GMV of the language/executive control network and behavioral executive performance. Highly expert bilinguals showed no relationship between performance levels and volumetric changes in their neural substrate, suggesting that they may be able to optimize executive performance even in the face of lower GMV. Below, we discuss these findings in more detail.

4.1 Effects of individual differences in bilingual experience factors on executive behavioral performance

In line with previous findings, our data indicate that increasing L2 proficiency is associated with better executive control performance in bilinguals. Similar results have been reported across different executive functions tasks including the ANT (Tao et al., 2011; Novitskiy et al., 2019), the Stroop task (Iluz-Cohen & Armon-Lotem, 2013; Singh & Mishra, 2012; 2013), Posner’s cueing paradigm (Mishra et al., 2012), and the Flanker task (Luk et al., 2011b; Sorge et al., 2017). As already discussed in the introduction, higher fluency in the L2 would cause bilinguals to experience higher cross-linguistic influence from their two linguistic systems (e.g., Van Hell & Dijkstra, 2002, Wu & Thierry, 2010). Thus, as a result of increased control demand, high-proficient bilinguals would develop a better ability to resolve conflict. At the same time, behavioral bilingualism studies have produced a number of inconsistent outcomes, even when deploying continuous measures of BEFs, with studies failing to firmly establish an effect of L2 proficiency on executive performance, either partially (i.e., in one of several executive tasks) or completely (e.g., Becker, Schubert, Strobach, Gallinat & Kühn, 2016; Dong & Xie, 2014; Rosselli, Ardila, Lalwani & Vélez-Uribe, 2016; Xie, 2018). Dong and Xie (2014), for instance, found no effect of varying L2 proficiency on the level of performance on the same Flanker task deployed in the present study. This and similar studies typically ascribe inconsistencies in the findings to the complex and largely compositional nature of both phenomena, bilingualism and executive control, leading to differing outcomes based on the choice of bilingual samples (e.g., balanced vs. unbalanced, young adults vs. seniors) as well as the operational definitions of executive control's sub-functions and the corresponding selection of experimental tasks (for a more detailed discussion, see Dong & Xie, 2014; Xie, 2018). The inconsistencies in the available behavioral results highlight that testing behavior may not be sufficient to get full insight into the effects of BEFs on executive control. We argue that a clearer picture can be achieved by investigating the consequences of differences in the bilingual experience for the neural substrate, as also exemplified by the brain morphometry results of the current investigation.

4.2 Effects of individual differences in bilingual experience factors on the language/executive control network

Individual differences in L2 exposure and proficiency predicted GMV in the bilateral ACC, with a similar trend (although only approaching statistical significance) emerging for the bilateral PFC. In the bilingual brain, the ACC is assumed to underlie cross-linguistic conflict resolution

(Abutalebi & Green, 2016), while the PFC is thought to support target response selection and non-target response inhibition during language control and language switching (Abutalebi & Green, 2016). In our investigation, L2 exposure and proficiency showed a crossed interaction, suggesting that a positive association between one predictor and GMV was present only when the value of the other predictor was low. For medium values of L2 exposure or proficiency, GMV volumes remained stable with the other BEF varying. When both predictors' values further increased, a decreasing trend in GMV emerged. This pattern suggests that bilingualism-related neuroplasticity could plateau at a certain stage, with neuroplastic changes ceasing and possibly even reverting with increasing bilingual experience. Our results are in line with Pliatsikas's Dynamic Restructuring Model (DRM; Pliatsikas, 2019). Based on a comprehensive review of the literature on neuroplasticity in bilinguals, L2 learners and simultaneous interpreters, the DRM attempts to formulate a time course for bilingualism-induced structural adaptation. Regarding cortical structures related to language control, as the ACC and PFC, the model predicts volumetric increases in the early stages of L2 acquisition, which would reflect increased effort imposed by controlling competing activations in the two languages. Such effects would be especially palpable in sequential bilinguals, i.e., individuals that started learning and using their L2 later than their native L1 (as our participants), with the volumetric increases expected to peak and subsequently disappear as language control becomes less effortful with increasing exposure and proficiency. The DRM explains the trajectory of bilingualism-induced neuroplasticity as pruning: bilinguals would develop extra connections to accommodate the increased language control effort, to subsequently eliminate the supernumerary connections once the more efficient ones are identified. This process would also account for the resilience to age-related neurodegeneration observed in aging bilinguals (see Gallo, Myachykov, Shtyrov & Abutalebi, 2020): the efficient connections surviving this pruning phase would also be the ones that survive age-related deterioration.

We argue that our results may reflect the mechanisms posited by the DRM, yet we refrain from making conclusive statements in this regard, as the present investigation includes structural, but no functional imaging measurements. Nevertheless, the aforementioned effects of BEF differences – and their interplay with GMV differences (as described below) – on modulating executive performance suggest that the consequences of bilingualism for cognition may evolve with bilingual experience: neurostructural increases might eventually be replaced by enhancements in functional efficiency. The results of the conjunct analysis of behavioral and neuroimaging data discussed hereafter might better inform us on the processes underlying this posited structure-to-function shift.

4.3 Relationships between individual differences in bilingual experience factors and grey matter volume in modulating executive control performance

As pointed out in the introduction section, by combining analyses at both behavioral and neural levels, we aimed to gain insight into the *qualitative*, in addition to the *quantitative*, changes induced by bilingual experience on cognition. Arguably our most interesting finding illuminates the role of BEFs in modulating the relationship between the executive neural substrate and the associated behavioral outcomes. For the bilateral caudate and the left PFC, indeed, at higher levels of bilingual experience (i.e., higher BI), executive performance was not affected by the differences in GMV; moreover, in line with our behavioral analysis's results, higher BI levels predicted overall better performance. Although the scope of the structural MRI analysis does not allow an "online" investigation of the underlying processes, the idea of a consistent correspondence between behavior and neural substrate naturally prompts the hypothesis of a functional enhancement that should underlie the results reported herewith, in the face of the absence of neurostructural causes. Functional neuroimaging studies comparing bilinguals and monolinguals support the interpretation of enhanced functional efficiency and flexibility of expert bilinguals' brain networks. Indeed, expert bilinguals have been reported to make a more efficient use of executive control areas showing lesser brain activation while outperforming monolingual peers at the behavioral level (e.g., Abutalebi et al., 2012). They were also shown to activate different/more extensive brain networks during executive task performance (i.e., enhanced network flexibility; see e.g., Luk, Anderson, Craik, Grady & Bialystok, 2010).

One might also speculate that the phenomenon reported here might constitute the source of the enhanced cognitive reserve observed in senior bilinguals (see Gallo et al., 2020). Our findings indicate that experienced bilinguals show optimal task performance in the face of GMV variations, with no effect of decreasing volumes on performance levels. This result, translated to the context of senescence, may suggest that expert senior bilinguals may be able to optimize their behavioral performance even in the face of age-related brain atrophy. Corroborating evidence in this direction comes from a study by Del Maschio et al. (2018), reporting a comparable pattern of results for expert bilingual, but not for monolingual aging individuals. Cognitive reserve is thought to arise from lifelong experiential factors that originate in early life, with evidence indicating an early presence of cognitive reserve in young adult populations (see Tucker & Stern, 2011). The findings reported here might therefore represent a sign of early cognitive reserve development, constituting the prerequisite to the protection against age-related cognitive decline observed in aging bilinguals, although further investigation is required to provide conclusive evidence in this direction.

Finally, the pattern illustrated in Figure 4 suggests that participants' performance converged at an optimal level in the face of increasing GMV and irrespectively of the level of BI. This result provides insight into the potential reasons behind inconsistencies in bilingual research. Indeed, it is well known that the existing evidence regarding bilingualism-related effects on executive performance is rather mixed, particularly in the young adult age group. Valian (2015) points to a possible cause of this inconsistency: young adults are usually engaged in many cognitively challenging activities (e.g., education, sports, social interactions, video-gaming, etc.) that could equal or surpass the cognitive challenges imposed by learning and controlling two languages. It might thus be difficult to observe cognitive consequences of bilingual experience in this age group without disentangling the contribution of bilingualism from other factors. Our results support this claim: when neuroplastic changes – induced by any experiential factor, not just bilingualism – reach sufficiently high levels, eventual consequences of bilingualism on executive performance may become unnoticeable.

On a more cautious side, it also has to be noted that some recent studies called into question the replicability of structural brain-behavior (SBB) relationships (e.g., Boekel et al., 2015; Masouleh, Eickhoff, Hoffstaedter, Genon, & Alzheimer's Disease Neuroimaging Initiative, 2019), particularly in the studies using healthy young populations and small sample sizes. However, when investigating SBB replicability in the context of ROI-based approaches, one of the few consistently replicable SBB relationships appeared to involve the executive component of the ANT (Boekel et al., 2015), that are identical to the Flanker Task within the ANT set-up utilized here, thus corroborating our results. Nevertheless, we invite the reader to bear these issues in mind when considering our results.

Another noteworthy point is that, potentially, it could be of great interest to directly compare BEFs trajectories between heterogeneous bilingual samples (such as the one here) and balanced bilinguals, to investigate possible qualitative differences in bilingualism-induced consequences for cognition related to different experiences of L2 acquisition. However, such investigations may be virtually impossible to implement in practice. Indeed, individuals in the standard balanced bilingual sample tend to have very similar levels of AoA (i.e. early simultaneous acquisition of L1 and L2) and proficiency at least, if not also of exposure as well. Such conditions would thus make it difficult to find enough variability in BEFs to be able to investigate the consequences of their variations on GMV or cognition. That said, this complication on its own should not prevent the use of behavioral measures, their individual and composite statistical indices and neuroanatomical structural variables in future studies of different bilingual samples.

5. Conclusion

Overall, the reported research shows that individual differences in BEFs play an important role in modulating the consequences of bilingualism for executive functioning, both at the behavioral and at the neural level. The present investigation highlights the importance of treating the multifaceted phenomenon of bilingualism as a continuous spectrum, departing from group comparisons to achieve better consistency in the results and shed light on a still hazy picture. Our investigation might provide us with better insight into the reasons behind inconsistencies emerging from research on the "classic" bilingual samples, i.e. those including early, balanced, high proficient bilinguals.

Finally, in the pursuit of understanding the interplay between bilingualism and cognition, we also advocate the need for the field to shift towards simultaneous investigations of the behavioral and neural consequences of bilingual experience, which may illuminate the qualitative, beyond quantitative, repercussions of dual language use on cognition.

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of neurolinguistics*, *20*, 242-275.

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

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*, *49*, 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*, 2126-2133.

Abutalebi, J., Guidi, L., Borsa, V., Canini, M., Della Rosa, P. A., Parris, B. A., & Weekes, B. S. (2015). Bilingualism provides a neural reserve for aging populations. *Neuropsychologia*, *69*, 201-210.

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and cognition*, *19*, 689-698.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear Mixed-Effects Models Using Eigen and S4. Available online at: http://cran.r-project.org/package=lme4

Becker, M., Schubert, T., Strobach, T., Gallinat, J., & Kühn, S. (2016). Simultaneous interpreters vs. professional multilingual controls: Group differences in cognitive control as well as brain structure and function. *Neuroimage*, *134*, 250-260.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in cognitive sciences*, *16*, 240-250.

Bialystok, E. (2016). The signal and the noise: Finding the pattern in human behavior. *Linguistic Approaches to Bilingualism*, *6,* 517-534.

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological bulletin*, *143*, 233.

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and cognitive processes*, *22*, 633-660.

Boekel, W., Wagenmakers, E. J., Belay, L., Verhagen, J., Brown, S., & Forstmann, B. U. (2015). A purely confirmatory replication study of structural brain-behavior correlations. *Cortex, 66,* 115-133.

Dahnke, R., Ziegler, G., & Gaser, C. (2012). Local adaptive segmentation. *HBM 2012*. http://www.neuro.unijena.de/hbm2012/HBM2012-Dahnke02.pdf

Del Maschio, N., Sulpizio, S., Gallo, F., Fedeli, D., Weekes, B. S., & Abutalebi, J. (2018). Neuroplasticity across the lifespan and aging effects in bilinguals and monolinguals. *Brain and cognition*, *125,* 118-126.

Del Maschio, N., Sulpizio, S., Toti, M., Caprioglio, C., Del Mauro, G., Fedeli, D., & Abutalebi, J. (2019). Second language use rather than second language knowledge relates to changes in white matter microstructure. *Journal of Cultural Cognitive Science*, 1-11.

DeLuca, V., Rothman, J., Bialystok, E., & Pliatsikas, C. (2019). Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proceedings of the National Academy of Sciences*, *116,* 7565-7574.

Diamond, A. (2013). Executive functions. *Annual review of psychology*, *64*, 135-168.

Dong, Y., & Xie, Z. (2014). Contributions of second language proficiency and interpreting experience to cognitive control differences among young adult bilinguals. *Journal of Cognitive Psychology*, *26*, 506-519.

ESS Round 8: European Social Survey Round 8 Data (2016). Data file edition 2.1. NSD - Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC. doi:10.21338/NSD-ESS8-2016.

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, 340-347.

Gallo, F., Myachykov, A., Shtyrov, Y., & Abutalebi, J. (2020). Cognitive and Brain Reserve in Bilinguals: Field Overview and Explanatory Mechanisms. *Journal of Cultural Cognitive Science,* doi:10.1007/s41809-020-00058-1. Published online by Springer Nature, May 13, 2020.

García-Pentón, L., Fernandez Garcia, Y., Costello, B., Duñabeitia, J. A., & Carreiras, M. (2016). The neuroanatomy of bilingualism: how to turn a hazy view into the full picture. *Language, Cognition and Neuroscience*, *31*, 303-327.

Gaser, C. (2020). *CAT, A Computational Anatomy Toolbox for SPM.* http://dbm.neuro.uni-jena.de/cat/

Gold, B. T., Johnson, N. F., & Powell, D. K. (2013). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia*, *51*, 2841-2846.

Gollan, T. H., Salmon, D. P., Montoya, R. I., & Galasko, D. R. (2011). Degree of bilingualism predicts age of diagnosis of Alzheimer's disease in low-education but not in highly educated Hispanics. *Neuropsychologia*, *49*, 3826-3830.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*, 515-530.

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

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

Guo, T., & Peng, D. (2006). Event-related potential evidence for parallel activation of two languages in bilingual speech production. *NeuroReport*, *17*, 1757-1760.

Hervais-Adelman, A., Egorova, N., & Golestani, N. (2018). Beyond bilingualism: multilingual experience correlates with caudate volume. *Brain Structure and Function*, *223*, 3495-3502.

Iluz-Cohen, P., & Armon-Lotem, S. (2013). Language proficiency and executive control in bilingual children. *Bilingualism: Language and Cognition*, *16*, 884-899.

Kaushanskaya, M., & Marian, V. (2007). Bilingual language processing and interference in bilinguals: Evidence from eye tracking and picture naming. *Language Learning*, *57*, 119–163.

Kroll, J.F., Bobb, S.C., & Hoshino, N. (2014). Two languages in mind: bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science*, *23*, 159-163.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, *1*, 377-394.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13).

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological bulletin*, *144*, 394-425.

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex*, *58*, 301-324.

Lüdecke D (2018). ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. *Journal of Open Source Software, 3*(26), 772. doi: 10.21105/joss.00772

Luk, G., Anderson, J. A., Craik, F. I., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and cognition*, *74*(3), 347-357.

Luk, G., Bialystok, E., Craik, F. I., & Grady, C. L. (2011a). Lifelong bilingualism maintains white matter integrity in older adults. *Journal of Neuroscience*, *31*, 16808-16813.

Luk, G., De Sa, E., & Bialystok, E. (2011b). Is there a relation between onset age of bilingualism and enhancement of cognitive control?. *Bilingualism: Language and cognition*, *14*, 588-595.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*, 605-621.

Macarthur foundation, 2007. *The MacArthur Research Network on Socioeconomic Status and Health.* https://macses.ucsf.edu/research/psychosocial/subjective.php#measurement

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*, 192-203.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*.

Masouleh, S. K., Eickhoff, S. B., Hoffstaedter, F., Genon, S., & Alzheimer's Disease Neuroimaging Initiative. (2019). Empirical examination of the replicability of associations between brain structure and psychological variables. Elife, 8, e43464.

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*, 757-757.

Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012). On the time course of exogenous cueing effects in bilinguals: higher proficiency in a second language is associated with more rapid endogenous disengagement. *The Quarterly journal of experimental psychology*, *65*, 1502-1510.

Mishra, R. K. (2015). Let's not forget about language proficiency and cultural variations while linking bilingualism to executive control. *Bilingualism: Language and Cognition*, *18,* 39-40.

Neurobehavioral Systems, Inc. (2020). *Presentation.* https://www.neurobs.com

Neuromorphometrics, Inc. *Building a Model of the Living Human Brain.* http://Neuromorphometrics.com/

Novitskiy, N., Myachykov, A., & Shtyrov, Y. (2019). Crosslinguistic interplay between semantics and phonology in late bilinguals: neurophysiological evidence. *Bilingualism: Language and Cognition, 22*, 209–227.

Novitskiy, N., Shtyrov, Y., & Myachykov, A. (2019). Conflict resolution ability in late bilinguals improves with increased second-language proficiency: ANT task evidence. *Frontiers in Psychology*, *10*, 2825.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*, 97-113.

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265-278.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., & Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human brain mapping*, *19*, 170-182.

Pliatsikas, C., Johnstone, T., & Marinis, T. (2014). Grey matter volume in the cerebellum is related to the processing of grammatical rules in a second language: a structural voxel-based morphometry study. *The Cerebellum*, *13*, 55-63.

Pliatsikas, C., DeLuca, V., Moschopoulou, E., & Saddy, J. D. (2017). Immersive bilingualism reshapes the core of the brain. *Brain Structure and Function*, *222*, 1785-1795.

Pliatsikas, C. (2019). Understanding structural plasticity in the bilingual brain: The Dynamic Restructuring Model. *Bilingualism: Language and Cognition*, 1-13.

R Core Team. (2014). *R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing*. Available online at: http://www.R-project.org/

Rajapakse, J. C., Giedd, J. N., & Rapoport, J. L. (1997). Statistical approach to segmentation of single-channel cerebral MR images. *IEEE transactions on medical imaging, 16*, 176-186.

Raven, J., Raven, J. C., & Court, J. H. (2000). Manual for Raven's progressive matrices and vocabulary scales. Section 3: The standard progressive matrices. Oxford, UK: Oxford Psychologists Press; San Antonio, TX: The Psychological Corporation.

Rosselli, M., Ardila, A., Lalwani, L. N., & Vélez-Uribe, I. (2016). The effect of language proficiency on executive functions in balanced and unbalanced Spanish–English bilinguals. *Bilingualism: Language and Cognition, 19*, 489-503.

Singh, N., & Mishra, R. K. (2012). Does language proficiency modulate oculomotor control? Evidence from Hindi–English bilinguals. *Bilingualism: Language and Cognition*, *15,* 771-781.

Singh, N., & Mishra, R. K. (2013). Second language proficiency modulates conflict-monitoring in an oculomotor Stroop task: Evidence from Hindi-English bilinguals. *Frontiers in psychology*, *4*, 322.

Sorge, G. B., Toplak, M. E., & Bialystok, E. (2017). Interactions between levels of attention ability and levels of bilingualism in children's executive functioning. *Developmental science*, *20*, e12408.

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*, 2015-2028.

Surrain, S., & Luk, G. (2019). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, *22*, 1-15.

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in psychology*, *2*, 123.

Tohka, J., Zijdenbos, A., & Evans, A. (2004). Fast and robust parameter estimation for statistical partial volume models in brain MRI. *Neuroimage, 23*, 84-97.

Tucker, A., & Stern, Y. (2011). Cognitive reserve in aging. *Current Alzheimer Research*, *8*, 354-360.

Valian, V. (2015). Bilingualism and cognition. *Bilingualism: Language and Cognition, 18*(1), 3-24.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic bulletin & review*, *9*, 780-789.

Wu, Y. J., & Thierry, G. (2010). Chinese–English bilinguals reading English hear Chinese. *Journal of Neuroscience*, *30*, 7646-7651.

Xie, Z. (2018). The Influence of Second Language (L2) Proficiency on Cognitive Control Among Young Adult Unbalanced Chinese-English Bilinguals. *Frontiers in psychology*, *9*, 412.

Yeung, N. (2014). Conflict monitoring and cognitive control. In: K. N. Ochsner and S. Kosslyn (Eds.) *The Oxford Handbook of Cognitive Neuroscience: Volume 2: The Cutting Edges*. Oxford-New York. Oxford University Press.

Footnotes

1. The term executive control (also referred to as executive functions or cognitive control; Diamond, 2013) refers to a multi-componential construct comprising a range of processes that support flexible, goal-directed behavior by representing task-relevant information in order to guide thought and action (Yeung, 2014).

Figure captions

Figure 1. 3D-rendered representation of regions of interest in the language control/executive control network. PFC = prefrontal cortex (note that this region is only partially represented for illustrative purposes); ACC = anterior cingulate cortex; IPL = inferior parietal lobule; CAU = caudate nucleus.

Figure 2. Interaction plot for the L2 proficiency*task condition interaction predicting Flanker's RTs (in ms). Increasing levels of L2 proficiency predict lower RT, i.e., better executive performance.

Figure 3. Interaction plot for the crossed L2 proficiency*L2 exposure interaction predicting left anterior cingulate cortex (LACC) GMV (in cm^3).

a. Higher L2 proficiency predicts increases in GMV only at lower levels of L2 exposure. For medium levels of L2 exposure, variations in L2 proficiency do not affect GMV. For high levels of L2 exposure, increases in L2 proficiency predict reductions in GMV.

b. Higher L2 exposure predicts increases in GMV only at lower levels of L2 proficiency. For medium levels of L2 proficiency, variations in L2 exposure do not affect GMV. For high levels of L2 proficiency, increases in L2 exposure predict reductions in GMV.

Comparable results emerged for the right anterior cingulate cortex, left prefrontal cortex and right prefrontal cortex. Normalized scores are reported for L2 proficiency and L2 exposure. Data was plotted with the sjPlot and ggeffects packages in R. Note that such packages, when a numeric vector is specified as a grouping structure, automatically select representative values for that vector (see Lüdecke, 2018 for more information).

Figure 4. Interaction plot for the bilingual index*task condition*GMV interaction predicting Flanker's RTs (in ms), for the left caudate nucleus (LCAU). Increases in GMV (in cm^3) predict lower incongruent RTs only at low scores of bilingual index. At increasing levels of bilingual index,

variations in GMV do not affect executive performance. Comparable results emerged for the right caudate nucleus and left prefrontal cortex. Normalized scores are reported for the bilingual index. Data was plotted with the sjPlot and ggeffects packages in R. Note that such packages, when a numeric vector is specified as a grouping structure, automatically select representative values for that vector (see Lüdecke, 2018 for more information).