

Verbal production dynamics and plasticity: functional contributions of language and executive control systems

Kshipra Gurunandan*¹, Manuel Carreiras^{1,2,3} and Pedro M. Paz-Alonso*^{1,2}

¹ BCBL Basque Center on Cognition, Brain and Language, 20009 Donostia-San Sebastian, Spain

² IKERBASQUE Basque Foundation for Science, 48013 Bilbao, Spain

³ Department of Basque Language and Communication, EHU/UPV, 48015 Bilbao, Spain

*Correspondence should be addressed to:

Kshipra Gurunandan or Pedro M. Paz-Alonso,

BCBL Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, 20009 Donostia-San Sebastian, Gipuzkoa, Spain.

Tel: +34 943-309-300, Fax: +34 943-309-052

E-mails: k.gurunandan@bcbl.eu, kepa.pazalonso@gmail.com

Running title: Language production dynamics and plasticity

Author Contributions: K.G., M.C., and P.M.P-A. designed research; K.G. and P.M.P-A. analysed data; K.G. wrote the paper, and P.M.P-A. and M.C. provided comments on the manuscript.

Abstract

Bilingual language production requires both language knowledge and language control in order to communicate in a target language. Learning or improving a language in adulthood is an increasingly common undertaking, and this has complex effects on the cognitive and neural processes underlying language production. The current fMRI experiment investigated the functional plasticity of verbal production in adult language learners, and examined the dynamics of word retrieval in order to dissociate the contributions of language knowledge and executive control. Thirty four adults who were either intermediate or advanced language learners, underwent MRI scanning while performing verbal fluency tasks in their native and new languages. A multi-pronged analytical approach revealed (i) time-varying contributions of language knowledge and executive control to verbal fluency performance, (ii) learning-related changes in the functional correlates of verbal fluency in both the native and new languages, (iii) no effect of learning on lateralisation, and (iv) greater functional coupling between language and language control regions with greater second language experience. Collectively, our results point to significant functional plasticity in adult language learners that impacts the neural correlates of production in both the native and new languages, and provide new insight into the widely-used verbal fluency task.

Keywords

bilingualism, fMRI, language learning, semantic retrieval, verbal fluency

1. Introduction

Learning or improving a new language in adulthood is an increasingly common and relevant skill. Bilingualism is known to effect functional and structural brain changes, but the bulk of the findings come from adults who acquired their second language in childhood. Adult language learners differ substantially from younger learners. Adults significantly outperform children in both syntactic and morphological learning in the early stages of language learning, and yet, adults have notoriously low ultimate achievement in new languages (Krashen et al. 1979; White and Genesee, 1996; Marinova-Todd et al. 2000). Furthermore, adult language learners are often able to achieve high levels of comprehension, but lag substantially in production (Walsh and Diller, 1979; Gershkoff-Stowe and Hahn, 2012). Consistent with these findings, neuroimaging studies of the early stages of adult language learning found changes in areas similar to those in younger learners (Mechelli et al. 2004; Barbeau et al. 2017). We also previously showed that language network lateralisation in adults displayed substantial learning-dependent plasticity in comprehension but not in production (Gurunandan et al. 2020). In the current study, we investigated the later stages (i.e. intermediate to advanced proficiency levels) of adult language learning, specifically, how learning a language in adulthood affects the cognitive and neural processes underlying language production.

Bilingual language production requires not only language knowledge but also language control. Language control allows bilinguals to selectively communicate in a target language while minimising interferences from the non-target language (Abutalebi et al. 2008; Branzi et al. 2020; Declerck et al. 2021). Knowledge and control both change with increasing second language proficiency: more proficient bilinguals have larger second language vocabulary as well as more refined language control. These multidimensional changes indicate that second language learning has substantial effect on the cognitive processes underlying language production and is likely to be associated with complex changes in the neural correlates of production. The goal of the current experiment was two-fold: (i) to investigate the functional plasticity of language production in the more proficient

stages of language learning, and (ii) to examine the neural dynamics of word retrieval and dissociate the contributions of language knowledge (i.e. vocabulary) and executive control.

In a typical word retrieval task, also known as the verbal fluency task, participants/patients are presented with a series of semantic categories such as “animals” or phonemic categories such as “words beginning with the letter A”, and asked to produce as many examples of each category as possible within a given time period, typically 60 seconds. This simple task has long been used in psycholinguistic research to study language production, as well as in clinical settings to evaluate executive functioning in healthy ageing and in a variety of disorders (Miller, 1984; Birn et al. 2010; Baciú et al. 2016). Verbal fluency, especially semantic category fluency, is thought to have components of both language and executive control, though the question of whether these components are differential and dissociable remains open (Shao et al. 2014; Whiteside et al. 2015; Aita et al. 2018). In the current study, we examined the idea that this hybrid nature of the verbal fluency task has a temporal component, i.e. language and executive control processes come into play at different stages of the task. Crowe (1998) studied the drop-off in the number of exemplars produced as the verbal fluency task progressed, and suggested that there is a store of high-frequency words which is more readily accessible, called the *topicon* (Smith and Claxton, 1972), and that once this is exhausted, the search is extended to a more extensive lexicon. In 2010, Luo and colleagues examined behavioural differences between monolinguals and low/high vocabulary bilinguals in verbal fluency performance and found complex interactions in the starting point and slopes of the time-course. Finally, in the context of time-varying contributions of the medial temporal lobe to semantic retrieval for categories that might be aided by episodic memory instances, it has been suggested that early responses are based on well-rehearsed prototypical knowledge while later responses rely more on open-ended strategies (Duff and Brown-Schmidt 2012; Sheldon and Moscovitch 2012). Here we propose that the beginning of the verbal fluency task is more dependent on lexical retrieval of frequent words from the *topicon*, and as the task progresses, more strategic executive processes come into play. The aforementioned behavioural studies performed correlations

between verbal fluency performance and cognitive scores in monolingual and early, high-proficiency bilingual populations, with mixed results, and in the current study, we conjectured that adult second language learning might magnify the differential contributions of language and executive processes and allow us to examine effects that might be too subtle in native or earlier acquired languages. The study thus offers important insights into bilingual language production and its effects on a task that is widely used in clinical settings.

We recruited adults enrolled in a language school in either the intermediate or the advanced level classes, and they performed verbal fluency tasks inside the scanner. If early word generation is associated with linguistic knowledge and later word generation with executive control, then we would expect performance in the first half of the verbal fluency task to be predicted by language proficiency, and performance in the second half, or perhaps the drop in performance in the second half, to be predicted by executive control. To comprehensively map the neural dynamics of verbal fluency and the effects of language learning, we used three analytical approaches: (i) functional activation of language and language control regions, (ii) lateralisation of activation in language and language control regions, and (iii) functional coupling between language and language control regions. First, we examined the time course of activation during the task. Previous studies of verbal fluency have found involvement of the frontal and temporal regions (e.g. Birn et al. 2010), though many studies have focused on the prefrontal cortex (e.g. Gaillard et al. 1999; Badre and Wagner, 2002; Costafreda et al. 2006; Wagner et al. 2014). These were static measures of activation, and we hypothesised that executive control regions would be more active as the task becomes progressively harder with each new response, and we expected to see greater group differences in the second half of the verbal fluency task once both groups had exhausted their topic. Second, we examined the lateralisation of activation during the verbal fluency task. Verbal fluency tasks have been previously used to localise language functioning in the brain and are known to elicit left-lateralised activation (Gaillard et al. 2003; Gurunandan et al. 2020), and here we were interested in the role of right hemisphere regions. In line with our previous findings, we did not expect to see any effect of L2

proficiency in lateralisation of language production, but hypothesised that as responses get more effortful in the later part of the task, right hemisphere contribution to the task might increase, leading to lower left-lateralisation in the second half of the task due to the proportionally lower contribution of the left hemisphere regions. Finally, we examined the role of L2 proficiency in functional coupling of the language and language control regions. In line with the idea that language usage enhances connectivity between language and language control regions, we found stronger connectivity during comprehension in more advanced L2 learners (Gurunandan et al. 2019), and this has also been found to be the case during picture naming in bimodal bilinguals (Li et al. 2015) and during the resting state, particularly in late bilinguals (Sulpizio et al. 2020). In the current study, we expected to observe stronger functional coupling in advanced learners who had more L2 experience, i.e. both proficiency and exposure, than the intermediate group.

2. Methods

2.1 Participants

The final study sample consisted of 31 right-handed native Spanish speakers (mean age = 45.19 \pm 10.64 years; 15 female) studying Basque in the same language school. Data from three additional participants was excluded due to either technical issues or excessive head motion during imaging. Participants were native to the Basque Country, Spain, and had grown up with Spanish (L1) at home and in school, with little exposure to Basque due to the prevailing sociopolitical situation. This situation changed in their adulthood and they were now living in Spanish-Basque bilingual environments, and were enrolled in Basque (L2) classes at either the A2 level (intermediate group, n = 17) or C1 level (advanced group, n = 14). The proficiency levels A2 and C1 correspond to those specified by the Common European Framework of Reference for Languages (CEFR), and participants had uniformly high performance in their respective classes (i.e. for their level). The study sample was thus controlled for both linguistic and extra-linguistic factors since all participants were native to the region and were learning a local language that differs substantially from their

native language in morphology and syntax, but has largely overlapping phonology and orthography. This ensured that motor learning of phonology and articulation were unlikely to be implicated (Berken et al. 2016). At the same time, beyond these aspects of experimental control, our hypotheses did not hinge on the linguistic distance between languages. Linguistic distance might entail different changes in regions such as the sensorimotor cortices, but we expect results in the language and control networks to be reasonably generalisable to different languages.

Language proficiency was further assessed using a picture-naming task in their two languages — an adaptation of the Boston Naming Test (Kaplan et al. 1983) controlled for cognates across L1 and L2 — and participants completed a language background questionnaire in which they indicated the percentage of their daily exposure outside the classroom to each language. The two groups were matched on age, gender, IQ, and Spanish proficiency (Table 1). Participants had limited knowledge of English or other languages, with little day-to-day exposure to them, and there was no difference between groups in this regard ($p = 0.91$). All participants had normal or corrected-to-normal vision, and no history of neurological or psychiatric disorders. In compliance with the ethical regulations established by the BCBL Ethics Committee and the guidelines of the Helsinki Declaration, all participants gave written informed consent prior to taking part in the experiment, and received monetary compensation for their participation.

Table 1: Participant demographics and linguistic scores

	Intermediate group		Advanced group		t-test
	Mean	SD	Mean	SD	<i>p</i> value
Age	44.47	11.14	46.07	10.34	0.682
Gender	8F, 9M		7F, 7M		
Fluid Intelligence	109.75	14.04	117.21	9.25	0.094
L1 proficiency	98.87	2.67	99.71	0.83	0.258
L2 proficiency	48	8.57	85.5	15.07	0.000
L1 Exposure (%)	85.87	15.79	70.54	23.95	0.063
L2 Exposure (%)	8.4	7.77	24.23	22.61	0.030

2.2 Task Design

Inside the scanner, participants performed an overt, paced semantic verbal fluency task in their two languages, i.e. Spanish and Basque, that was designed to allow investigation of the time-course of behaviour and activation. Several studies have demonstrated that covert or silent speech, though popular in fMRI paradigms, does not activate the same networks as overt speech (Barch et al. 1999; Huang et al. 2001; Gracco et al. 2005; Shuster and Lemieux, 2005; Sörös et al. 2006; Christoffels et al. 2007; Kleber et al. 2007). Careful head fixation has been found to minimise motion artefacts from overt speech (Heim et al. 2006), and to control articulatory motion and activation throughout the task, participants were instructed to move as little as possible while speaking and to respond overtly with “pass” if they had no response. Paced responses allowed control over timing of responses, and the number of exemplars produced during paced and free recall tasks inside the scanner have been found to be similar (Basho et al. 2007). Finally, we used semantic categories since this is most directly relevant to natural language production (Levelt et al. 1999).

The in-scanner task had a block design with two runs per language, each run containing eight semantic categories and eight control blocks. To avoid language-switching, the languages were separated and the order of languages was counterbalanced across participants. In each run, participants fixated on a white cross in the middle of a black screen, and semantic category cues (e.g. fruits, animals, clothes) were presented on the screen. Each cue was displayed eight times over the course of 30 seconds, and participants were instructed to respond overtly to each cue with an exemplar (e.g. fruits: apple, pear, banana, etc.). Each cue required a novel response, or failing this, an overt response saying “pass” in order to control articulatory motion and activation throughout the task. In the control condition, participants repeated the word presented on the screen. Repetitions, inflections of the same word and erroneous responses were removed, and fluency was scored as percent valid answers averaged across the sixteen categories in each language.

Outside the scanner, participants also performed a standard attention network test (ANT; Fan et al. 2002), and the executive attention component was used as a measure of executive control. The

task was carried out in an insulated behavioural testing cabin. Participants fixated on a black cross in the middle of a white screen, and in each trial, a single line of five arrows was presented either above or below the fixation cross. Participants were instructed to indicate the direction of the central arrow by pressing either the left or the right button as quickly and as accurately as possible. All participants performed practice trials in which they were given feedback (an emoticon appeared on the screen after each response), and there was no feedback in the experimental trials.

2.3 MRI Data Collection

Whole-brain MRI data was acquired using a 3-T Siemens Magnetom Trio whole-body MRI scanner and a 32-channel head coil at the Basque Center on Cognition, Brain and Language (BCBL). Padded headphones were used to dampen background scanner noise and enable clear transmission of the auditory stimuli. Participants viewed the print stimuli on a screen via a mirror mounted on the head coil. To limit head movement, the head coil was padded with foam and participants were asked to remain as still as possible.

Functional MRI was acquired in the course of four separate runs using a gradient-echo echo-planar pulse sequence with the following parameters: TR 3000 ms, TE 25 ms, 43 axial slices with a 3.0 x 3.0 x 3.0 mm voxel resolution, 10% inter-slice gap, flip angle (FA) = 90°, field of view (FoV) = 192 mm, 64 x 64 matrix. 240 volumes were collected for each of the functional runs. Prior to each scan, four volumes were discarded to allow for T1-equilibration effects. Structural T1-weighted images were acquired with a MPRAGE sequence with TR = 2530 ms, TE = 2.97 ms, inversion time = 1100 ms, FA = 7°, FoV = 256 x 256 mm, 176 slices and voxel size = 1 mm³.

2.4 Data Analyses

For all analyses, there were three main factors of interest: (i) between-subjects factor Group (intermediate or advanced), (ii) within-subject factor Language (L1, L2), and (iii) within-subject factor Time (A, B) i.e. the task was partitioned into two parts, A and B, where A comprised the first half of the task, and B comprised the second half of the task.

Behavioural performance

Each response during the in-scanner verbal fluency task was scored and averaged across semantic cues to track the mean performance during the course of each task block. Performance was next averaged across the first and second halves of the task (Time A and B respectively), and a mixed-model ANOVA was conducted, with Group (intermediate, advanced) as a between-subjects factor, and Language (L1, L2) and Time (A, B) as within-subject factors. We also ran Bayesian ANOVAs and t-tests to calculate the default Bayes Factor (BF; Jeffreys 1961; Rouder et al. 2012) as a measure of the evidence against the null hypothesis. BF quantifies the relative predictive performance of two rival hypotheses, for e.g. $BF = 5$ indicates that the data are 5 times more likely under the alternative hypothesis than under the null hypothesis. While one of the advantages of BF is that it offers a quantification of the evidence on a continuous scale, a BF of 1-3 is often considered weak evidence for the alternative hypothesis, $BF > 3$ as substantial evidence, and $BF > 10$ as strong evidence for the alternative hypothesis, while $BF < 0.33$ is typically considered substantial evidence for the null hypothesis. The ANT reaction times (RT) were used to calculate a measure of executive control: mean RT of all congruent flanking conditions, summed across cue types, subtracted from the mean RT of incongruent flanking conditions (Fan et al. 2002).

To test our hypotheses about the hybrid nature of the verbal fluency task and separate the roles of *a priori* independent variables language proficiency and executive control, multiple regression analyses were run. Since task performance was at ceiling in L1, we focused on task performance in L2. Hence L2 performance was modelled as the response variable, with L2 proficiency and executive control (ANT) as explanatory variables, and L1 proficiency and fluid intelligence as control variables. The four independent variables were tested for independence using pair-wise correlation, and then entered into multiple linear regression models. We tested four dependent variables: (i) task performance in the first half of the task, i.e. part A, (ii) task performance in the second half of the task, i.e. part B, (iii) overall task performance, i.e. A + B, and (iv) drop in task performance from first to second halves of the task, i.e. A - B.

L2 Task Performance in first half (A) ~ L1 proficiency + Fluid Intelligence + L2 proficiency + Executive Control
L2 Task Performance in second half (B) ~ L1 proficiency + Fluid Intelligence + L2 proficiency + Executive Control
Overall L2 Task Performance (A + B) ~ L1 proficiency + Fluid Intelligence + L2 proficiency + Executive Control
Drop in L2 Task Performance (A - B) ~ L1 proficiency + Fluid Intelligence + L2 proficiency + Executive Control

Finally, correlations were plotted between task performance and its significant predictors.

fMRI pre-processing

Standard SPM12 (Penny et al. 2011) preprocessing routines and analysis methods were employed. Images were first corrected for differences in timing of slice acquisition and then realigned to the first and mean volumes using rigid-body registration. Each subject's functional volumes were spatially smoothed with a 4-mm full-width half-maximum (FWHM) Gaussian kernel. Next, motion parameters obtained from realignment were used to inform a volume repair procedure (ArtRepair version 5b; Mazaika et al. 2009) that identified bad volumes on the basis of scan-to-scan movement (>1 mm) and signal fluctuations in global intensity (>1.3%), and corrected bad volumes via interpolation between the nearest non-repaired scans. Data from two subjects requiring more than 20% of their volumes to be repaired were discarded. The number of corrected volumes was similar between groups ($p=0.43$). After volume repair, functional volumes were co-registered to the T1 images using 12-parameter affine transformation and spatially normalised to the MNI space by applying non-linear transforms estimated by deforming the MNI template to each individual's structural volume. During normalisation, the volumes were sampled to 3-mm cubic voxels. The resulting volumes were then spatially smoothed with a 7-mm FWHM Gaussian kernel. Due to the quadratic relation between separate smoothing operations, the total smoothing applied to the functional data was approximately equivalent to smoothing with an 8-mm FWHM Gaussian kernel. Finally, time series were temporally filtered to eliminate contamination from slow frequency drift (high-pass filter with cut-off period of 128 seconds).

fMRI analyses

Statistical analyses were performed on individual subject data using the general linear model (GLM). fMRI time series data were modelled by a series of impulses convolved with a canonical haemodynamic response function (HRF). Six motion parameters for translation (x, y, z) and rotation (yaw, pitch, roll) were included as covariates of non-interest in the GLM. Each trial was modelled as an epoch of 15 seconds each, time-locked to either the beginning or the middle of the presentation of each block. The remaining functions were used as covariates in the GLM, along with a basic set of cosine functions that high-pass filtered the data, and a covariate for session effects. The least-squares parameter estimates of the height of the best-fitting canonical HRF for each study condition were used in pairwise contrasts. Brain coordinates throughout the text, as well as in tables and figures, are reported in MNI atlas space.

All subsequent analyses were performed with *a priori* neuroanatomical regions of interest (ROI) chosen from standard neuroanatomical models of language (e.g. Hagoort, 2013) and language control (Abutalebi and Green, 2007). We chose six language ROIs: inferior frontal gyrus (IFG) *pars orbitalis*, IFG *pars triangularis*, IFG *pars opercularis*, inferior parietal cortex (IPC), posterior superior temporal gyrus (STG), and ventral occipito-temporal cortex (vOTC), using the AAL atlas (Tzourio-Mazoyer et al. 2002). For language control, we chose two ROIs: anterior cingulate cortex (ACC) and dorsolateral prefrontal cortex (dLPFC), extracting the caudal ACC and rostral middle frontal gyrus respectively from the Desikan-Killiany atlas (Desikan et al. 2006) since these regions were less refined in the AAL atlas. By using anatomically-defined ROIs from widely-tested models and atlases, we aimed to maximise the generalisability and replicability of our results, and avoid using *a posteriori* functionally-defined ROIs that would be far more specific to our experimental design. Additionally, since using unbiased *a priori* ROIs was particularly important for the lateralisation analysis (described below), using anatomically-defined ROIs for all analyses allowed for consistency across the different analyses.

ROI analysis

Parameter estimates (% signal change) of the eight left hemisphere ROIs were acquired for the contrast Semantics>Rest in each language condition, and ran mixed-model ANOVAs with Group (intermediate, advanced) as a between-subjects factor, and Language (L1, L2) and Time (A, B) as within-subject factors. Since we were testing eight individual ROIs, p -values were corrected for multiple comparisons using false discovery rate (FDR), and then simple effects tests were run on interaction effects that were significant at corrected $p < 0.05$. We also calculated BF values to quantify the evidence against the null hypothesis. To further examine the course of functional activation during the task, the time course analysis of activation was plotted for the same ROIs. BOLD signal time-series were extracted from each ROI by averaging time-series across all voxels in each ROI. Condition-wise task blocks were each modelled as 30s windows of activity and their time-series were averaged together to construct mean time courses for each language. These condition-averaged time courses were then averaged across functional runs, and subsequently plotted.

Lateralisation

Laterality indices of activation for the eight ROIs were calculated in each language condition, and submitted to mixed model ANOVAs, with Group (intermediate, advanced) as a between-subjects factor, and Language (L1, L2) and Time (A, B) as within-subject factors. Since we were testing eight ROIs, p -values were corrected for multiple comparisons using FDR. We also calculated BF values to quantify the evidence against the null hypothesis. Lateralisation is measured by using bilateral ROI masks to calculate the relative contributions of left and right hemisphere regions, with the following equation:

$$\textit{Laterality Index} = \frac{\textit{Activation in left hemisphere} - \textit{Activation in right hemisphere}}{\textit{Total activation}}$$

This returns a value between -1 and $+1$, with negative values indicating greater activation in the right hemisphere region compared to the left homologue, and positive values indicating greater activation in the left hemisphere region compared to the right homologue. In line with the latest

recommendations (Bradshaw et al. 2017), we used a threshold-independent method to calculate the index using the LI-Toolbox (Wilke and Lidzba 2007) and *a priori* anatomically-defined regions. The LI-Toolbox was used with the recommended default settings, i.e. a bootstrapped activation threshold, with inclusive bilateral masks extracted from the AAL and Desikan-Killiany atlases.

Functional Connectivity

To test changes in functional coupling between regions as the task progressed, functional connectivity was assessed using the beta-series correlation method (Rissman et al. 2004) implemented in SPM12 with custom MATLAB scripts. The canonical HRF in SPM was fitted to each occurrence of each condition and the resulting parameter estimates (beta values) were sorted according to the study conditions to produce a condition-specific beta series for each voxel. Pairwise connectivity was calculated between selected ROIs in the bilateral language networks and control regions for each participant and condition. Since the correlation coefficient ranges from -1 to $+1$, an arc-hyperbolic tangent transform (Fisher 1915) was applied to these beta-series correlation values to make its null hypothesis sampling distribution approach that of the normal distribution. The normally distributed Fisher's Z values for each of the eight left hemisphere ROIs were submitted to mixed model ANOVAs, with Group (intermediate, advanced) as a between-subjects factor, with Language (L1, L2) and Time (A, B) as within-subject factors, and p -values were corrected for multiple comparisons using FDR. We also calculated BF values to quantify the evidence against the null hypothesis.

Data availability

Data are available at <https://www.bcbl.eu/DataSharing/CerebCor2022Gurunandan/>

3. Results

3.1 In-scanner Behavioural Performance

The mixed-model ANOVA revealed a Group x Language interaction in participants' fluency scores ($F(1,23)=36.75$, $q=0.000003$, $BF=32780.45\pm 3.09\%$). Post-hoc simple-effect analyses showed that the intermediate and advanced learners exhibited no difference in fluency in their L1 ($t(24.99)=0.45$, $p=0.654$, $BF=0.386\pm 0\%$), but a significant difference in L2 ($t(21.6)=5.92$, $p<0.001$, $BF=2739.5\pm 0\%$), with the advanced group showing significantly higher L2 fluency than the intermediate group. Both groups displayed significantly higher performance in L1 than in L2 (intermediate group: $t(13)=16.01$, $p<0.001$, $BF=14335775.1\pm 0\%$, and advanced group: $t(10)=4.90$, $p<0.001$, $BF=60.31\pm 0\%$) (Figure 1A). Finally, there was a significant main effect of Time, with task performance significantly higher in the first half of the task A compared to the second half of the task B ($F(1,23)=87.8$, $q<0.001$, $BF=71.34\pm 0.76\%$) (Figure 1B). The time course of behavioural performance was plotted for each Group and Language (Figure 1C).

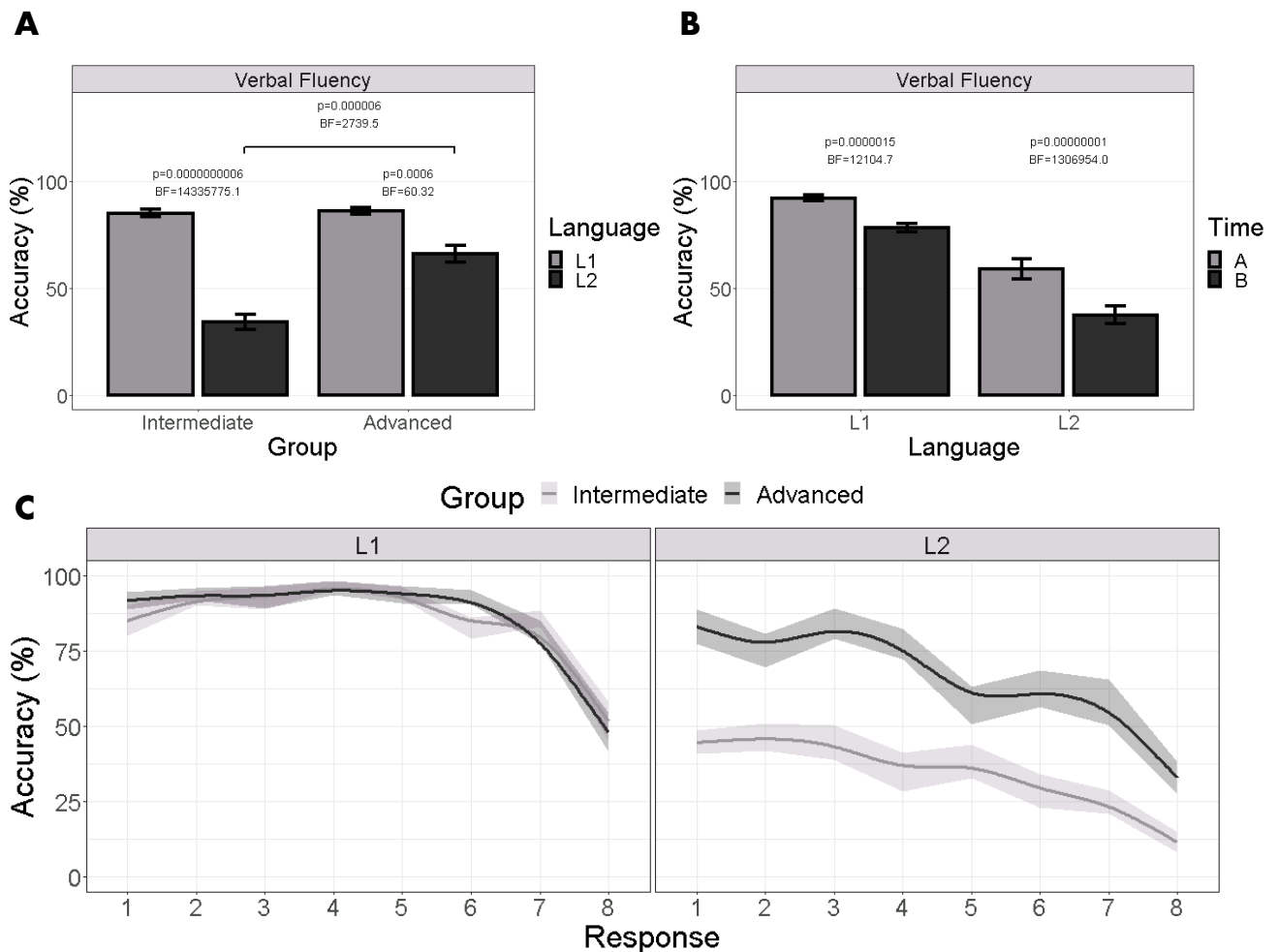


Figure 1. Behavioural results of the verbal fluency task inside the scanner revealed (A) Group x Language interaction in percent correct responses, (B) Main effect of Time in percent correct responses. (C) Time course of behavioural performance for each Group and Language was plotted for illustrative purposes. Error bars represent standard deviation. BF represents the default Bayes Factor.

Multiple regression analyses were used to test if L2 proficiency and executive control significantly and dissociably predicted participants' L2 verbal fluency performance, when controlling for L1 proficiency and fluid intelligence. The four independent factors did not display any significant pairwise correlations. L2 performance in both the first and second halves of the task was significantly predicted by L2 proficiency alone. The same effect was seen in overall L2 performance, i.e. A + B, which was significantly predicted by L2 proficiency ($\beta=0.76$, $p<0.001$), but not by executive control ($p=0.658$). The model explained 72.07% of the variance in L2 performance ($R^2=0.72$, adjusted

$R^2=0.66$, $F(4,17)=10.97$, $p<0.001$). L2 performance was significantly correlated with L2 proficiency ($r=0.83$, $p<0.001$; Figure 2A), but not with executive control ($r=0.13$, $p=0.5$).

The drop in L2 performance from first to second halves of the task, i.e. A - B, was significantly predicted by executive control ($\beta=292.9$, $p=0.002$), while L2 proficiency did not show a significant effect ($\beta=0.2$, $p=0.06$). The model explained 54.63% of the variance in the drop in L2 performance ($R^2=0.55$, adjusted $R^2=0.44$, $F(4,17)=5.12$, $p=0.006$). The drop in L2 performance was significantly correlated with executive control ($r=0.57$, $p=0.002$), but not with L2 proficiency ($r=0.23$, $p=0.27$). For easier interpretation of the scatter-plot, we plotted the reciprocal of the ANT RT value such that greater scores indicate greater executive control, and plotted B - A such that higher scores indicate more sustained performance throughout the task (Figure 2B).

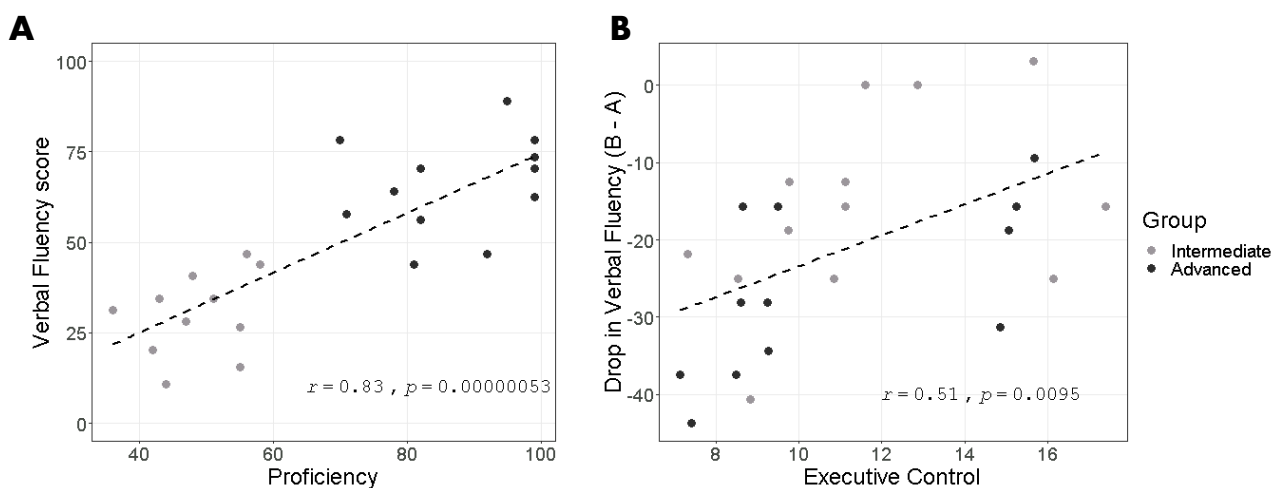


Figure 2. Pearson’s correlations revealed (A) positive correlation between L2 proficiency and L2 VF performance, and (B) positive correlation between executive control and sustained L2 VF performance (i.e. drop in performance from first to second half of the task).

3.2 ROI Analysis

The overall pattern of recruitment of regions in L1 and L2 was exactly the same (Figure 3A). Mixed model ANOVAs of the activation in each ROI revealed a main effect of Language in the ACC, dlPFC and IFG (*pars orbitalis*, *pars triangularis*, and *pars opercularis*), with significantly higher activation for L1 than for L2 (Figure 3A). We found significant Group x Time interactions in the dlPFC, IFG *pars triangularis*, and IFG *pars opercularis*, with significantly higher activation in the

advanced than the intermediate group in only the second half of the task (Figure 3b). The ROI time courses of activation were plotted for each Group and Language (Figure 3C).

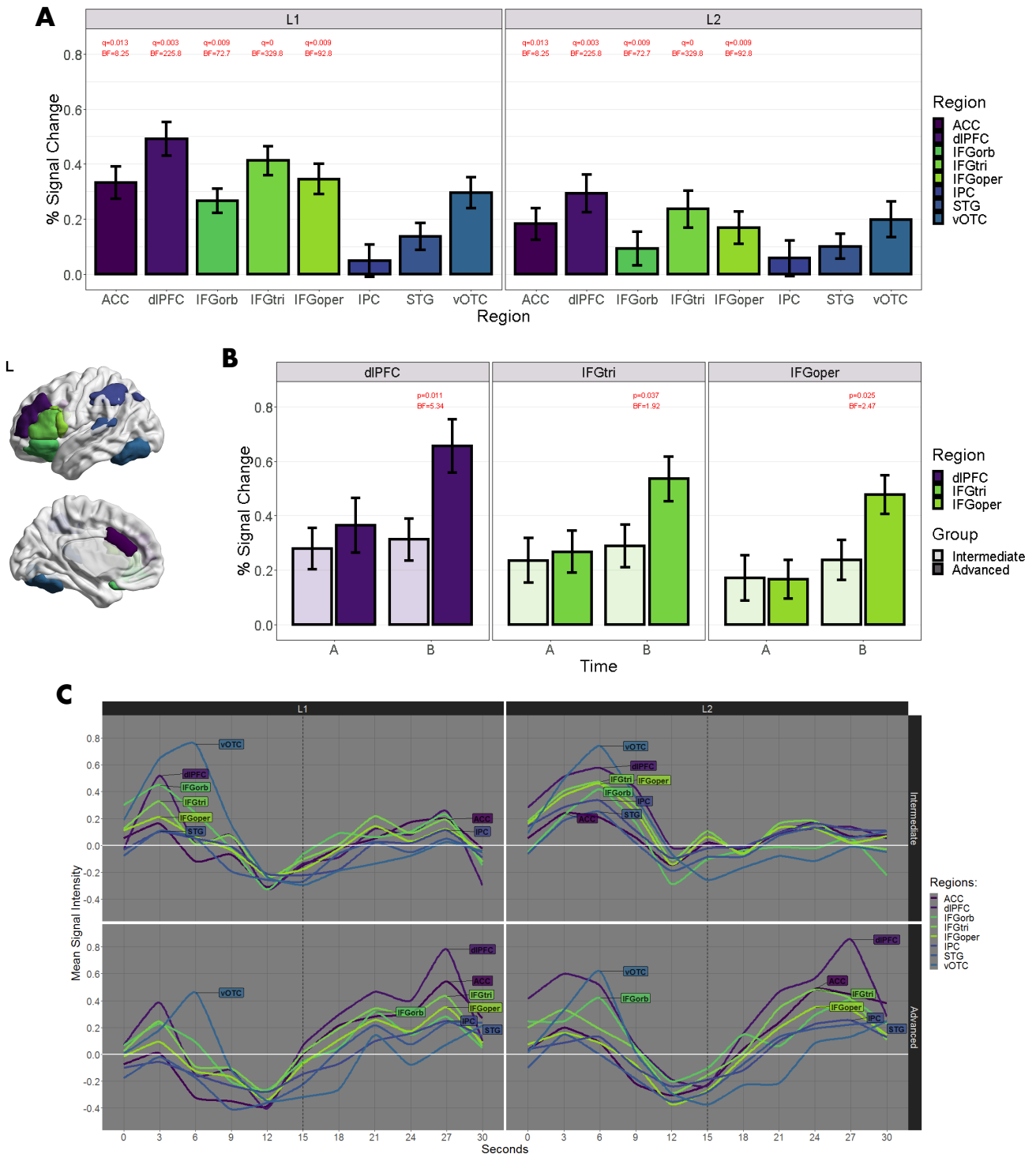


Figure 3. ROI analysis of the left hemisphere *a priori* ROIs during the verbal fluency task revealed (A) similar pattern of activation in L1 and L2, and a main effect of Language in the control and IFG regions, and (B) Group x Part interaction in control regions and Broca's Area. (C) Time course of ROI activation for each Group and Language was plotted for illustrative purposes. Language control regions=purple, IFG=green, and posterior regions=blue. Error bars represent standard error, q indicates FDR-corrected p-values, and BF indicates default Bayes Factor.

3.3 Lateralisation

Verbal production in the language network regions was left-lateralised. Mixed model ANOVAs of the laterality indices in each of the eight bilateral ROIs revealed no effect of proficiency in any of the regions ($q > 0.65$; $BF < 0.8$, $BF_{\text{error}} < 0.01$). We found a main effect of Language in the IFG *pars triangularis* and STG, with significantly greater left lateralisation in L1 than in L2 (Figure 4A). We also found a main effect of Time in only the language regions, i.e. IFG *pars orbitalis*, *triangularis*, and *opercularis*, IPL and STG, with significantly greater left lateralisation in the first half of the task, i.e. Time A (Figure 4B). Control regions ACC and dlPFC did not display this effect of Time ($q > 0.17$; $BF < 0.5$, $BF_{\text{error}} < 0.02$).

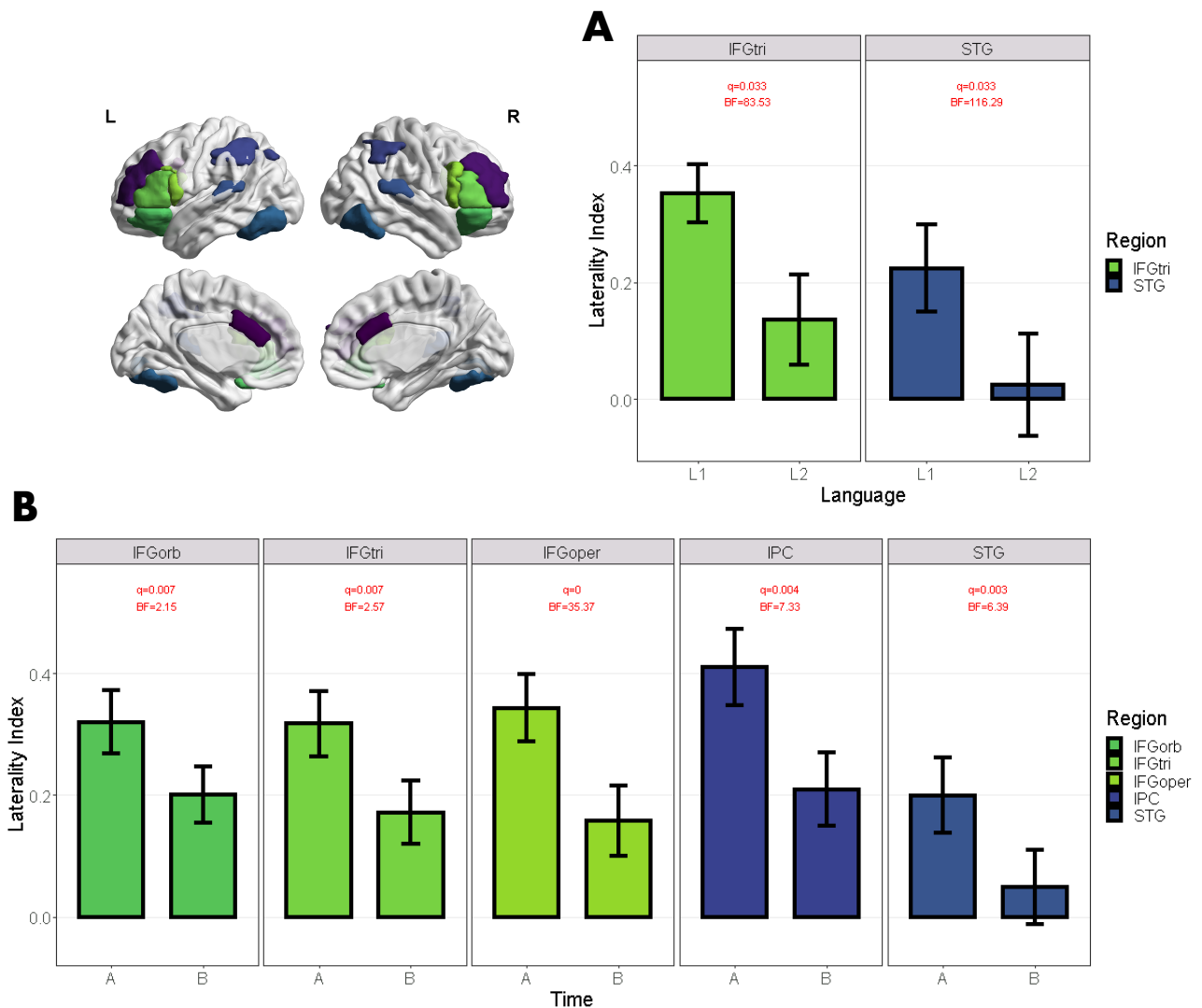
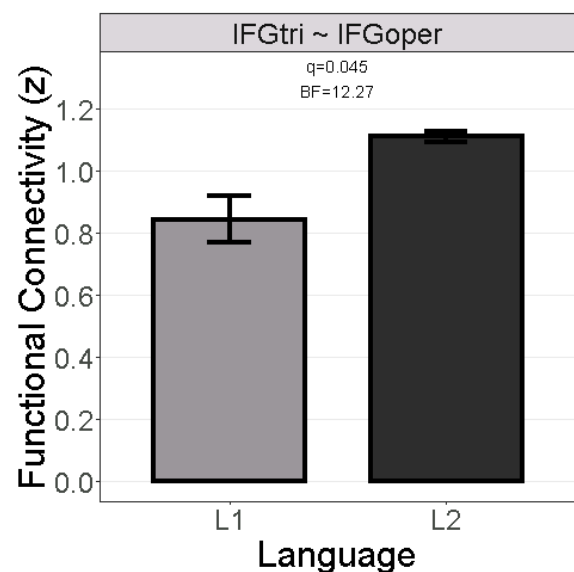
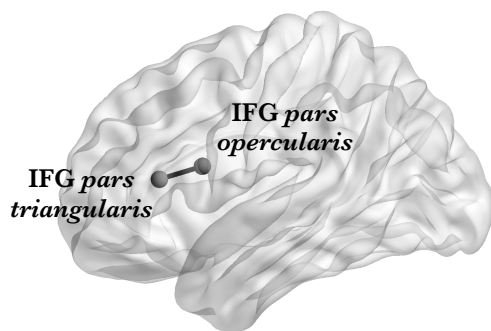


Figure 4. Laterality analysis of the bilateral *a priori* ROIs during the verbal fluency task revealed (A) main effect of Language and (B) main effect of Time in the language regions. Language control regions=purple, IFG=green, and posterior regions=blue. Error bars represent standard error, q indicates FDR-corrected p-values, and BF indicates default Bayes Factor.

3.4 Functional Connectivity

Mixed model ANOVAs of the functional connectivity between the eight left hemisphere ROIs revealed a main effect of Language in the connectivity between the IFG *pars triangularis* and *opercularis*, with significantly higher connectivity in L2 than in L1 (Figure 5A). We also found a main effect of Group in the connectivity between the ACC and IFG *pars triangularis*, with significantly higher connectivity in the advanced group compared to the intermediate group (Figure 5B). There was no significant effect of Time in any ROI pair ($q > 0.22$).

A



B

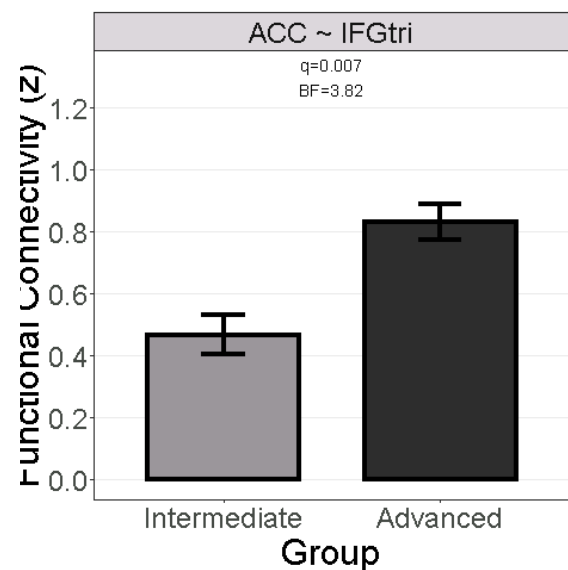
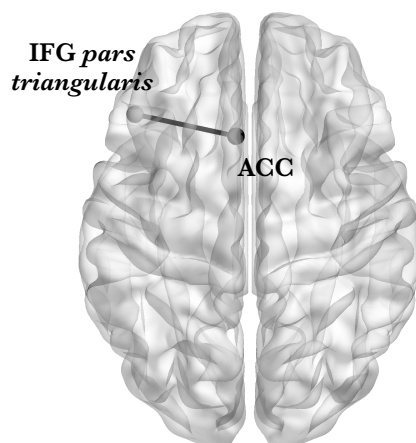


Figure 5. Functional connectivity analysis of the left hemisphere a priori ROIs during the verbal fluency task revealed (A) main effect of Language in functional coupling in the IFG and (B) main effect of Group in functional coupling of language and language control regions. Error bars represent standard error, q indicates FDR-corrected p-values, and BF indicates default Bayes Factor.

4. Discussion

The current study investigated functional differences between intermediate and advanced language learners with two goals: (i) to investigate the effect of second language learning on the neural substrates of language production, and (ii) to examine the neural dynamics of verbal fluency and dissociate the contributions of language knowledge (i.e. vocabulary) and executive control. To the best of our knowledge, this is the first fMRI study to investigate the later stages (i.e. more proficient levels) of ecological adult language learning and the first to examine the functional dynamics of bilingual verbal fluency. The interaction between language and proficiency groups allowed us to test various hypotheses about the functional dynamics of verbal fluency, and to examine effects in a recently acquired language that might be too small to be detected in native or early languages. To ensure sufficient L2 proficiency differences, we used a cross-sectional design and studied two groups of adult language learners from intermediate and advanced level classes at the same language school. All participants were studying a local language with very different morphology and syntax, but similar phonology and the same writing system and a transparent orthography as their native language, thus controlling for both extra-linguistic factors such as immigration or foreignness as well as linguistic factors such as motor learning of phonology and articulation. Inside the MRI scanner, participants performed semantic verbal fluency tasks in their native and new languages. Participants were presented with semantic category words on the screen (eight times in a row per category), and instructed to respond aloud with a new exemplar every time they saw the category word on the screen. If they could not think of a new exemplar, they were instructed to say “pass” in the appropriate language. The overt, paced task was designed to allow investigation of its time course, and all analyses were performed by partitioning the task into two parts. We first confirmed our hypothesis about the time course of different cognitive processes during the verbal fluency task, showing that language knowledge was associated with overall verbal fluency performance and that greater executive control was associated with a lower drop in performance as the task became progressively more difficult. Our fMRI analytical approach

included ROI analyses of functional activation, lateralisation analyses, and analyses of functional connectivity. We found (i) significant learning-related changes in the functional correlates of verbal fluency in the native and new language, (ii) no significant learning-related changes in lateralisation, and (iii) increased functional coupling between language and language control regions with greater L2 experience. Collectively, our results point to significant functional plasticity in the later stages of language learning in adults that affect both the native and new languages, and support our hypotheses about the cognitive and neural dynamics of verbal fluency.

Verbal fluency is predicted by language proficiency and executive control

The two groups were equated in their L1 proficiency and performed similarly in the L1 verbal fluency task, but displayed significant differences in their L2 proficiency and verbal fluency task performance. Since L1 proficiency and verbal fluency were at ceiling, we did not expect proficiency or executive control to be predictive of verbal fluency performance, and this was indeed the case. In L2, however, in which both groups had significantly lower performance, we found that verbal fluency was significantly associated with language proficiency but not executive control, while greater executive control was associated with a smaller drop in performance from the first to second halves of the task which was not predicted by language proficiency. This provides support to the hypothesis that the beginning of the semantic verbal fluency task is more dependent on language knowledge, and as the task progresses, more strategic executive processes come into play. Previous studies of verbal fluency mostly agree that semantic verbal fluency is associated with language knowledge (e.g. Friesen et al. 2015; Escobar et al. 2018), but dissociating it from executive control has proven to be much harder, with far less consensus on whether semantic verbal fluency tasks should be considered language tasks (e.g. Luo et al. 2010) or executive control tasks (e.g. Aita et al. 2018). In the current study, the association between L2 proficiency and verbal fluency was driven by the large difference in L2 proficiency between the two groups; however, there was no difference between groups in executive control as measured by the ANT, and executive control was found to be significantly predictive of the trajectory of performance regardless of language knowledge. The lack

of consensus in the field indicates that the dissociable effects of language knowledge and executive control are quite small in healthy adults in their native and early languages, but it is critical to understand and account for the interaction between these two factors in verbal fluency performance due to the widespread use of such tasks in clinical settings. Further studies with bilingual patients should be carried out to confirm our findings in the relevant clinical settings.

L2 proficiency alters the neural dynamics of verbal fluency in both L1 and L2

The pattern of recruitment of the language and language control regions was identical in both languages, indicating that similar neural processes underlay verbal fluency in both languages, though the control regions, i.e. ACC and dlPFC, and the IFG displayed greater activation for L1 than for L2. We found significant effects of L2 proficiency in the functional dynamics of verbal fluency. In both languages, we found that in the first half of the task, there was no significant difference between intermediate and advanced groups in any region, while in the second half of the task, the advanced group displayed significantly greater recruitment of the dlPFC and the IFG *pars triangularis* and *pars opercularis*, i.e. Broca's Area. The time course of activation further illustrated the difference in dynamics between the two groups, and the presence of this effect in both languages despite L1 being equated in both groups and L2 being significantly different between them indicates a complex interaction between language knowledge and language control during word retrieval. While greater L2 proficiency is associated with a larger L2 lexicon, the consequently lower functional frequency of both L1 and L2 words in the mental lexicon increases reliance on executive search strategies (Costa, 2005; Michael and Golan, 2005). Thus, acquiring a new language and becoming proficient in it fundamentally alters the cognitive and neural dynamics of word retrieval in both native and new languages.

L2 proficiency does not affect left lateralisation of language production

The lateralisation of activation in the language and control regions during the verbal fluency task exhibited no effects of proficiency. In our previous work, we had found that lateralisation of the overall language network during verbal production does not change with increasing L2 proficiency

(Gurunandan et al. 2019). Here we examined the neural dynamics of lateralisation during language production in individual language and language control regions, and once again found that there was no effect of L2 proficiency on lateralisation of language production. However, we found that lateralisation was modulated by task difficulty, i.e. left lateralisation of language regions decreased with increasing task difficulty, indicating greater recruitment of right hemisphere homologues. This was seen in two instances: (i) the IFG *pars triangularis* and posterior STG were significantly more left-lateralised in L1 than in L2 in both groups throughout the task, and (ii) the language regions displayed significantly higher left-lateralisation of activation at the beginning of the task compared to the end of the task. Previous studies have found similar effects, with greater right hemisphere activation in more difficult language tasks (Buckner et al. 1995; Schnur et al. 2009). A few previous studies have hypothesised a non-language-specific role of the right hemisphere regions (e.g. Basho et al. 2007; Vigneau et al. 2011; Geranmayeh et al. 2014), but in the current experiment, the drop in left-lateralisation was specific to the language regions, supporting the idea that the involvement of the right hemisphere regions at the end of the task is specific to language and not executive processes. Further, the decrease in left-lateralisation in the second half of the task coincided with the increased activation of left-hemisphere regions in the second half of the task that was found in the activation analysis, indicating that lower left-lateralisation was not caused by lower left hemisphere activation, but instead by greater right hemisphere activation of language regions.

Greater L2 experience enhances functional coupling of language and control regions

Pairwise functional connectivity analyses revealed a main effect of L2 proficiency on functional coupling of the ACC with IFG *pars triangularis*. We found that advanced L2 learners exhibited significantly stronger coupling than did the intermediate L2 learners, indicating that coordination between these language and control regions was significantly higher during language production across languages. We found no effect of task progression, indicating that functional connectivity was relatively constant throughout the verbal fluency task and primarily dependent on language experience. Finally, we found that functional coupling between the IFG *pars triangularis* and *pars*

opercularis was higher in L2 than in L1, indicating greater connectivity within the language network for L2 processing. Overall, the finding that functional coupling of language and language control regions is higher in advanced L2 learners is consistent with previous findings such as bimodal bilinguals showing greater connectivity between ACC and language regions during picture-naming compared to monolinguals (Li et al. 2015), advanced learners displaying higher connectivity between control and language regions during an L2 semantic judgement task (Gurunandan et al. 2019), and bilinguals with greater language experience exhibiting greater resting state connectivity between control and language networks (Sulpizio et al. 2020). In the current study, L2 proficiency and daily L2 exposure were not independent, and thus based on previous studies, we suggest that greater L2 experience, i.e. proficiency and exposure, results in enhanced functional coupling of language and control regions.

Limitations

First, the current study has a modest sample size. We were thus careful to focus on specific hypotheses and to examine the data with a multi-pronged approach by including Bayesian statistics, interpreting each result within the context of corroborating results from diverse analyses, and staying within the scope of our results. Second, we used *a priori* neuroanatomical ROIs that, while unbiased, are larger than functional ROIs, and this may have the drawback of being less precise. Third, in the interest of limiting the number of ROIs and maintaining consistency with our previous work, we did not examine sub-cortical regions implicated in language control (Bradley et al. 2013; Hervais-Adelman et al. 2015). Finally, while participants' L2 proficiency was carefully characterised and controlled to the best of our ability, we cannot rule out contributions of language learning talent and active learning effort that may have contributed to the higher level of proficiency in the advanced group.

Conclusion

In conclusion, we found that across languages, intermediate and advanced language learners displayed significant differences in (i) the functional dynamics of word retrieval, and (ii) functional

connectivity between language and language control regions. Thus language learners well into adulthood display functional plasticity of language and language control networks. Importantly, these effects are not specific to the new language, but affect the native language as well, indicating that learning a new language fundamentally alters language production processes in all languages. We also found that language proficiency and executive control play dissociable roles in semantic verbal fluency, and together with the finding that these are affected by second language learning, it is critical to further understand and account for the effects of bilingualism on verbal fluency performance in clinical settings.

Acknowledgements

Kshipra Gurunandan received support from “la Caixa” Foundation (ID 100010434) through the fellowship LCF/BQ/DI17/11620005 and the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 713673. Manuel Carreiras was supported by project APCIN-2015-061-MultiLateral funded by the Spanish Ministry of Economy and Competitiveness [FLAG-ERA JTC 2015], a grant RTI2018-093547-B-I00 from the Spanish Ministry of Science and Innovation, and a grant from “la Caixa” Banking Foundation under the project code LCF/PR/HR19/52160002. Pedro Paz-Alonso was supported by grants from the Spanish Ministry of Science and Innovation [RYC-2014-15440; PGC2018-093408-B-I00], Neuroscience projects from the Fundación Tatiana Pérez de Guzmán el Bueno, and Basque Government [PIBA-2021-1-0003]. The research was also supported by funding from the Basque Government through the BERC 2022-2025 program and by the Spanish State Research Agency through BCBL Severo Ochoa excellence accreditation CEX2020-001010-S.

References

- Abutalebi J, Annoni JM, Zimine I, Pegna AJ, Seghier ML, Lee-Jahnke H, Lazeyras F, Cappa SF, Khateb A. 2008. Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb Cortex*. 18(7):1496-505.
- Abutalebi J, Green D. 2007. Bilingual language production: The neurocognition of language representation and control. *J Neurolinguistics*. 20(3):242-75.

- Aita SL, Beach JD, Taylor SE, Borgogna NC, Harrell MN, Hill BD. 2018. Executive, language, or both? An examination of the construct validity of verbal fluency measures. *Appl Neuropsychol Adult*. 26(5):441-451.
- Baciu M, Boudiaf N, Cousin E, Perrone-Bertolotti M, Pichat C, Fournet N, Chainay H, Lamalle L, Krainik A. 2016. Functional MRI evidence for the decline of word retrieval and generation during normal aging. *Age*. 38:3.
- Badre D, Wagner AD. 2002. Semantic retrieval, mnemonic control, and prefrontal cortex. *Behav Cogn Neurosci Rev*. 1(3):206-18.
- Barbeau EB, Chai XJ, Chen JK, Soles J, Berken J, Baum S, Watkins KE, Klein D. 2017. The role of the left inferior parietal lobule in second language learning: An intensive language training fMRI study. *Neuropsychologia*. 98:169-76.
- Barch DM, Sabb FW, Carter CS, Braver TS, Noll DC, Cohen JD. 1999. Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. *Neuroimage*. 10(6):642-57.
- Basho S, Palmer ED, Rubio MA, Wulfeck B, Müller RA. 2007. Effects of generation mode in fMRI adaptations of semantic fluency: paced production and overt speech. *Neuropsychologia*. 45(8):1697-706.
- Berken JA, Gracco VL, Chen JK, Klein D. 2016. The timing of language learning shapes brain structure associated with articulation. *Brain Struct Funct*. 221:3591–3600.
- Birn RM, Kenworthy L, Case L, Caravella R, Jones TB, Bandettini PA, Martin A. 2010. Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. *Neuroimage*. 49(1):1099-107.
- Bradley KA, King KE, Hernandez AE. 2013. Language experience differentiates prefrontal and subcortical activation of the cognitive control network in novel word learning. *Neuroimage*. 67:101-10.
- Bradshaw AR, Bishop DV, Woodhead ZV. 2017. Methodological considerations in assessment of language lateralisation with fMRI: a systematic review. *PeerJ*. 5:e3557.
- Branzi FM, Martin CD, Carreiras M, Paz-Alonso PM. 2020. Functional connectivity reveals dissociable ventrolateral prefrontal mechanisms for the control of multilingual word retrieval. *Hum Brain Mapp*. 41(1):80-94.
- Buckner RL, Petersen SE, Ojemann JG, Miezin FM, Squire LR, Raichle ME. 1995. Functional anatomical studies of explicit and implicit memory retrieval tasks. *J Neurosci*. 15(1):12-29.
- Christoffels IK, Formisano E, Schiller NO. 2007. Neural correlates of verbal feedback processing: an fMRI study employing overt speech. *Hum Brain Mapp*. 28(9):868-79.
- Costa A. 2005. Lexical access in bilingual production. In Kroll JF, De Groot AM, editors. *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press. p 308–325.
- Costafreda SG, Fu CH, Lee L, Everitt B, Brammer MJ, David AS. 2006. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Hum Brain Mapp*. 27(10):799-810.
- Crowe SF. 1998. Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. *J Clin Exp Neuropsychol*. 20(3):391-401.
- Declerck M, Meade G, Midgley KJ, Holcomb PJ, Roelofs A, Emmorey K. 2021. On the Connection Between Language Control and Executive Control—An ERP Study. *Neurobiol Lang (Camb)*. 4:1-9.
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS. 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 31(3):968-80.
- Duff MC, Brown-Schmidt S. 2012. The hippocampus and the flexible use and processing of language. *Front Hum Neurosci*. 6:69.
- Escobar GP, Kalashnikova M, Escudero P. 2018. Vocabulary matters! The relationship between verbal fluency and measures of inhibitory control in monolingual and bilingual children. *J Exp Child Psychol*. 170:177-89.

- Fan J, McCandliss BD, Sommer T, Raz A, Posner MI. 2002. Testing the efficiency and independence of attentional networks. *J Cogn Neurosci*. 14(3):340-7.
- Fisher RA. 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*. 10(4):507-521.
- Friesen DC, Luo L, Luk G, Bialystok E. 2015. Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Lang Cogn Neurosci*. 30(3):238-50.
- Gaillard WD, Hertz-Pannier L, Mott SH, Barnett AS, LeBihan D, Theodore WH. 1999. Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. *Neurology*. 54(1):180.
- Gaillard WD, Sachs BC, Whitnah JR, Ahmad Z, Balsamo LM, Petrella JR, Braniecki SH, McKinney CM, Hunter K, Xu B, Grandin CB. 2003. Developmental aspects of language processing: fMRI of verbal fluency in children and adults. *Hum Brain Mapp*. 18(3):176-85.
- Geranmayeh F, Wise RJ, Mehta A, Leech R. 2014. Overlapping networks engaged during spoken language production and its cognitive control. *J Neurosci*. 34(26):8728-40.
- Gershkoff-Stowe L, Hahn ER. 2012. Word comprehension and production asymmetries in children and adults. *J Exp Child Psychol*. 114(4):489-509.
- Gracco VL, Tremblay P, Pike B. 2005. Imaging speech production using fMRI. *Neuroimage*. 26(1):294-301.
- Gurunandan K, Arnaez-Telleria J, Carreiras M, Paz-Alonso PM. 2020. Converging Evidence for Differential Specialization and Plasticity of Language Systems. *J Neurosci*. 40(50):9715-24.
- Gurunandan K, Carreiras M, Paz-Alonso PM. 2019. Functional plasticity associated with language learning in adults. *Neuroimage*. 201:116040.
- Hagoort P. 2013. MUC (memory, unification, control) and beyond. *Front Psychol*. 4:416.
- Heim S, Amunts K, Mohlberg H, Wilms M, Friederici AD. 2006. Head motion during overt language production in functional magnetic resonance imaging. *Neuroreport*. 17(6):579-82.
- Hervais-Adelman A, Moser-Mercer B, Michel CM, Golestani N. 2015. fMRI of simultaneous interpretation reveals the neural basis of extreme language control. *Cereb Cortex*. 25(12):4727-39.
- Huang J, Carr TH, Cao Y. 2002. Comparing cortical activations for silent and overt speech using event-related fMRI. *Hum Brain Mapp*. 15(1):39-53.
- Jeffreys H. 1961. *The theory of probability*. New York: Oxford University Press.
- Kaplan EF, Goodglass H, Weintraub S. 1983. *The Boston naming test*. Philadelphia: Lea & Febiger.
- Kleber B, Birbaumer N, Veit R, Trevorrow T, Lotze M. 2007. Overt and imagined singing of an Italian aria. *Neuroimage*. 36(3):889-900.
- Krashen SD, Long MA, Scarcella RC. 1979. Age, rate and eventual attainment in second language acquisition. *TESOL quarterly*. 1:573-82.
- Lee MD, Wagenmakers EJ. 2014. *Bayesian cognitive modeling: A practical course*. New York: Cambridge university press.
- Levelt WJ, Roelofs A, Meyer AS. 1999. A theory of lexical access in speech production. *Behav Brain Sci*. 22(1):1-38.
- Li L, Abutalebi J, Zou L, Yan X, Liu L, Feng X, Wang R, Guo T, Ding G. 2015. Bilingualism alters brain functional connectivity between “control” regions and “language” regions: Evidence from bimodal bilinguals. *Neuropsychologia*. 71:236-47.
- Luo L, Luk G, Bialystok E. 2010. Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*. 114(1):29-41.
- Marinova-Todd SH, Marshall DB, Snow CE. 2000. Three misconceptions about age and L2 learning. *TESOL quarterly*. 34(1):9-34.
- Mazaika PK, Hoeft F, Glover GH, Reiss AL. 2009. Methods and software for fMRI analysis of clinical subjects. *Neuroimage*. 47(Suppl 1):S58.

- Mechelli A, Crinion JT, Noppeney U, O'Doherty J, Ashburner J, Frackowiak RS, Price CJ. 2004. Structural plasticity in the bilingual brain. *Nature*. 431(7010):757-757.
- Michael EB, Gollan TH. 2005. Being and becoming bilingual. In Kroll JF, De Groot AM, editors. *Handbook of bilingualism: Psycholinguistic approaches*. New York: Oxford University Press. p 389-407.
- Miller E. 1984. Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. *Br J Clin Psychol*. 23(1):53-7.
- Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE. 2011. *Statistical parametric mapping: the analysis of functional brain images*. London: Elsevier.
- Rissman J, Gazzaley A, D'Esposito M. 2004. Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage*. 23(2):752-63.
- Rouder JN, Morey RD, Speckman PL, Province JM. 2012. Default Bayes factors for ANOVA designs. *J Math Psychol*. 56(5):356-74.
- Schnur TT, Schwartz MF, Kimberg DY, Hirshorn E, Coslett HB, Thompson-Schill SL. 2009. Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of Broca's area. *Proc Natl Acad Sci U S A*. 106(1):322-327.
- Shao Z, Janse E, Visser K, Meyer AS. 2014. What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Front Psychol*. 5:772.
- Sheldon S, Moscovitch M. 2012. The nature and time-course of medial temporal lobe contributions to semantic retrieval: An fMRI study on verbal fluency. *Hippocampus*. 22(6):1451-66.
- Shuster LI, Lemieux SK. 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain Lang*. 93(1):20-31.
- Smith PT, Claxton GL. 1972. *Lexical search and phonemic organisation in memory*. Experimental Psychology Society, London.
- Sörös P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ, Stuss DT. 2006. Clustered functional MRI of overt speech production. *Neuroimage*. 32(1):376-87.
- Sulpizio S, Del Maschio N, Del Mauro G, Fedeli D, Abutalebi J. 2020. Bilingualism as a gradient measure modulates functional connectivity of language and control networks. *Neuroimage*. 205:116306.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 15(1):273-89.
- Vigneau M, Beaucousin V, Hervé PY, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N. 2011. What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage*. 54(1):577-93.
- Wagner S, Sebastian A, Lieb K, Tüscher O, Tadić A. 2014. A coordinate-based ALE functional MRI meta-analysis of brain activation during verbal fluency tasks in healthy control subjects. *BMC Neurosci*. 15(1):1-3.
- Walsh TM, Diller KC. 1979. Neurolinguistic considerations on the optimum age for second language learning. In *Annual Meeting of the Berkeley Linguistics Society*. 5:510-524.
- White L, Genesee F. 1996. How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Lang Res*. 12(3):233-65.
- Whiteside DM, Kealey T, Semla M, Luu H, Rice L, Basso MR, Roper B. 2015. Verbal fluency: Language or executive function measure?. *Appl Neuropsychol Adult*. 23(1):29-34.
- Wilke M, Lidzba K. 2007. LI-tool: a new toolbox to assess lateralization in functional MR-data. *J Neurosci Methods*. 163(1):128-36.