The neural basis of free language choice in bilingual speakers:
Disentangling language choice and language execution

Carlo Reverberi [1,2], Anna K. Kuhlen [3,4,5], Shima Seyed-Allaei [1,2,6], R. Stefan Greulich [3,4,7], Albert Costa [8,9], Jubin Abutalebi [10], John-Dylan Haynes [3,4,5]

1 Department of Psychology, Università Milano - Bicocca, Milan, Italy

2 NeuroMI - Milan Center for Neuroscience, Milan, Italy

3 Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health (BIH)

4 Bernstein Center for Computational Neuroscience, Berlin Center for Advanced Neuroimaging, Department of Neurology, and Excellence Cluster NeuroCure, Berlin, Germany

5 Humboldt-Universität zu Berlin, Berlin School of Mind and Brain and Institute of Psychology, Berlin,Germany

6 School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

7 Neurobiology Laboratory, German Primate Center GmbH, Göttingen, Germany

8 Center of Brain and Cognition (CBC), Universitat Pompeu Fabra, Barcelona, Spain

9 Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

10 Centre for Neurolinguistics and Psycholinguistics, University Vita-Salute San Raffaele & Scientific Institute San Raffaele, Milan, Italy

ABSTRACT

For everyday communication bilingual speakers need to face the complex task of rapidly choosing the most appropriate language given the context, maintaining this choice over the current communicative act, and shielding lexical selection from competing alternatives from non-target languages. Yet, speech production of bilinguals is typically flawless and fluent. Most of the studies available to date constrain speakers' language choice by cueing the target language and conflate language choice with language use. This left largely unexplored the neural mechanisms underlying free language choice, i.e., the voluntary situation of choosing the language to speak. In this study, we used fMRI and Multivariate Pattern Analysis to identify brain regions encoding the target language when bilinguals are free to choose in which language to name pictures. We found that the medial prefrontal cortex encoded the chosen language prior to speaking. By contrast, during language use, language control recruited a wider brain network including the left inferior frontal lobe, the basal ganglia, and the angular and inferior parietal gyrus bilaterally. None of these regions were involved in language choice. We argue that the control processes involved in language choice are different from those involved in language use. Furthermore, our findings confirm that the medial prefrontal cortex is a domain-general region critical for free choice and that bilingual language choice relies on domain general processes.

Keywords: Bilingualism; Language; MVPA; Free choice; Cognitive control; Time resolved fMRI; Speech; Intention; Lexicon; Naming

INTRODUCTION

The choice of which language to speak is an integral aspect of language production in bilinguals. It involves activatory and inhibitory mechanisms allowing bilinguals to select the proper word in the target language while withholding its translation equivalent in other languages (Abutalebi & Green, 2007; Green, 1998). Language control functions must guarantee the choice of the target language and protection from interferences from the alternative languages (Green & Abutalebi, 2013).

The cognitive processes underlying language control can be divided in different components such as 1) making the choice to speak in a given language, 2) the maintenance of this choice during verbalization, 3) the selection of the target word and the possible inhibition of words from the non-target language, 4) the monitoring of the output for potential intrusions like viable candidate words in the other language, and finally 5) language disengagement and engagement, i.e., stop speaking in one language and start speaking another language (Costa, Miozzo, & Caramazza, 1999; Green & Abutalebi, 2013; Kroll, Bobb, & Wodniecka, 2006; Reverberi et al., 2015). These components may rely on different neural (and thus likely cognitive) substrates. Indeed, in a recent study we showed that different neural structures are involved in changing and maintaining a language cued choice, and in executing such choice (Reverberi et al., 2015).

Language control and language choice have been extensively studied, typically using cued picture-naming tasks (Costa & Santesteban, 2004; Meuter & Allport, 1999). In such tasks, a cue indicating the language to be used for naming is presented before, together, or even after the picture to be named. If the language to be used changes between the consecutive pictures to be named, a switch trial will result. It has been shown that switching languages has a cost, so that switching trials lead to slower reaction times than non-switch trials. (Meuter & Allport, 1999). However, cued switching paradigms provide limited information on how language choice occurs since these paradigms override the voluntary action of deciding which language to produce. Hence, to better understand how language selection and control work we need to assess how different contextual situations may affect these processes (Gollan et al. 2009; 2014). Indeed, it is likely these processes are affected by whether or not the speaker is free to decide which language to use (Gollan & Ferreira, 2009; Kleinman & Gollan, 2016). For example, Gollan and collaborators showed that whenever speakers are not forced into overriding their natural naming preferences and instead can select their language freely, the costs associated with keeping both languages accessible are relatively small or non-existent (Kleinman & Gollan, 2016).

Not only the switching costs may be reduced when bilinguals can choose a language freely. Also, more generally, making the choice to speak a certain language, the maintenance thereof**,** and the execution of this choice may rely on distinct processes in a free choice setting. On a neural level, evidence from almost two dozens of neuroimaging studies shows that language control is achieved through a network tightly related to domain general executive control (for reviews Abutalebi & Green, 2007, 2016). A meta-analysis on the available literature on language switching highlighted that this network mainly involves the prefrontal cortex, the inferior parietal lobules, the anterior cingulate cortex (ACC) and the basal ganglia (Luk et al., 2012). Critically, however, neuroimaging studies investigating language control are based almost exclusively on cued paradigms (but see Zhang et al., 2015), thus not informing us about settings in which bilinguals freely choose a language. Furthermore, almost all these studies collapse the choice and the execution phases of the language selection process, implicitly assuming that they rely on the same cognitive and neural mechanisms, while this does not seem to be the case (Reverberi et al., 2015).

In the current study we explore the neural structures involved in language choice and language execution when bilinguals are free to choose in which language pictures are to be named. We do so by asking German - English bilinguals to freely choose which language to use to name line drawings, while their brain activity was scanned by means of fMRI. Language choice and language execution were temporally separated by a pure maintenance time delay. Multivariate pattern analysis (Haynes & Rees, 2006; Kamitani & Tong, 2005; Kriegeskorte, Goebel, & Bandettini, 2006; Norman, Polyn, Detre, & Haxby, 2006) was applied to identify regions encoding the chosen language during language choice and maintenance.

METHODS

Participants

Native German speakers with very good knowledge of English were recruited with flyers and emails. Volunteers who indicated that English was their second mother tongue (early bilinguals) or who reported high fluency in a third language were not recruited. All other volunteers were invited for a first experimental session, in which language proficiency was assessed. Assessment of language proficiency was based on a German version of the Language Experience and Proficiency Questionnaire (LEAP-Q Marian, Blumenfeld, & Kaushanskaya, 2007), and subjects' performance in a picture naming task that used a procedure similar to the main experiment (see result section for more details). None of the subjects chose more than 65% of the times the preferred language in this phase. The final subject group (n = 17) was selected on the basis of these language proficiency tests and consisted of highly proficient late bilinguals. Two subjects were excluded because they asked to abort scanning. The remaining 15 subjects with full scanning data (age mean 23,13 (SD=2.83), 10 female) were right handed. All subjects had normal or corrected to normal vision. No subject reported a neurological or psychiatric history, or a language or learning impairment. The subjects came from the same sample used for our previous study (Reverberi et al., 2015), and were scanned on average two months after being scanned for the first experiment. The experiment was approved by the ethics committee of the department of psychology of the Humboldt Universität zu Berlin. Subjects received monetary reward for their participation.

Stimuli

During the main experiment subjects were asked to name 120 black and white line drawings of objects extracted from the database of the International Picture Naming Project (IPNP) (Szekely et al., 2004). The database provides object-naming norms for 520 drawings of common objects for six languages, including German and American English. We excluded objects associated to homonyms in German and English. Furthermore, to ensure that there were not systematic differences in word length, we also exclude objects associated to words with different number of syllables in the two languages. From the remaining drawings we selected 120 objects. In our selection we maximized the word agreement (i.e. the proportion of subjects using a specific word to name an object) by applying a cut-off at 1.6 to the average H statistics of each item (Snodgrass & Vanderwart, 1980) for English and German. Secondarily, we minimized the expected naming latency difference between languages. Thus, in the selected set of 120 words, the mean of word agreement between English (0.465 SD: 0.477) and German (0,496 SD: 0.481) is not significantly different (p = 0.596, Mann-Whitney U). The same is true for the expected naming latency (English 956 ms SD: 203 for English and 1001 SD:214 ms for German; p = 0.0762, Mann-Whitney U). Word frequency was not considered for selection, but it is also not significantly different between languages (p = 0.1247, Mann-Whitney U).

Experimental Procedure

The fMRI experiment was divided into six runs, each comprising 40 trials. Each fMRI run lasted on average 9.26 minutes. During each experimental trial (Figure 1), subjects were required to name a line drawing of an object using the language of their choice. Each trial consisted of (1) the language choice phase, (2) the maintenance phase, (3) the picture naming phase, followed (4) by a distraction task.

Trials began with the presentation of a blank screen with a green fixation cross (400 ms) followed by a white cross (1600 ms). Afterward, the fixation cross turned red (400 ms). Between the green and the red cross subjects could freely choose between German or English (language choice phase lasting 2 s). Subjects were told to choose spontaneously a language in each trial without considering their choice in previous trials, and without following a predefined order. It was stressed that there was no right or wrong choice. Once a choice was made, subjects were asked not to change it. Upon red cross presentation subjects had to maintain the chosen language and prepare to respond (maintenance phase lasting 1 to 6 s). The red cross was followed by a white cross lasting 5600 ms in regular trials, and 600 ms in catch trials (see below). Finally, a line-drawing was presented for 3 s (picture naming phase), within which subjects had to generate a verbal response. The response was recorded by an MRI compatible microphone. Subjects were asked to minimize jaw movements involved in speaking, in order to try to avoid movement related artifacts in the EPI images.

After the picture naming, a distraction task filled the time before the beginning of the next trial. This task was introduced to prevent subjects from already making a decision on which language to use in the upcoming trial. In the distraction task, a rotating box with a missing side was presented. Subjects had to press a button whenever the missing side of the box pointed downwards. The distraction task lasted between 1 s to 6 equally distributed, with an overall run average of 3.5 s. Eight catch trials with a shorter maintenance phase (1 s overall) were randomly introduced into each run to ensure that subjects were deciding which language to use at the time required and that they were keeping their decision active for task execution.

Before fMRI scanning subjects underwent a short behavioral training during which subjects could familiarize themselves with the naming task. For this purpose we presented ten line-drawings that were not used in the main experiment. The training phase lasted about 5 minutes. Besides training with the experimental task, subjects could also practice uttering answers with minimal jaw movements. Matlab and the Cogent toolbox (http://www.vislab.ucl.ac.uk/cogent.php) were used to present the stimuli.

Image acquisition

Functional MRI scanning was performed at the Berlin Center for Advanced Neuroimaging (BCAN) using a 3 tesla Siemens Trio scanner equipped with a 12-channel head coil. In each of the six scanning sessions, an average of 290 T2*-weighted gradient-echo echo-planar images (EPI) containing 33 slices (3 mm thick) separated by a gap of 0.75 mm were acquired. Imaging parameters were as follows: repetition time (TR), 2000 ms; echo time (TE), 30 ms; flip angle, 78; matrix size, 64 x 64; and a field of view (FOV), 192 x 192 mm; thus yielding an in-plane voxel resolution of 3 mm^2, resulting in a voxel size of 3 x 3 x 3.75 mm. T1-weighted structural dataset was also collected. The parameters were as follows: TR, 1900 ms; TE, 2.52 ms; matrix size, 256 x 256; FOV, 256 mm; 192 slices (1 mm thick); flip angle, 9°. Finally we acquired field maps to correct for distortions of the magnetic field. All parameters were kept identical to EPI images but the TR = 400 ms, the TE = 5.19 ms / 7.65 ms, and the flip angle 60°.

The microphone FOMRI-II from Optoacoustics LTD was used for recording the verbal responses of the subjects during fMRI acquisition. The environment in the scanner during EPI acquisition is characterized by loud and high frequency acoustic noise. This microphone was developed for safe usage in an MRI environment. The microphone also featured a noise cancellation system for reducing the load high-frequency noise in the scanner environment during EPI acquisition. All audio recordings were further cleaned in a post-acquisition phase with a noise cancellation algorithm and a low pass filter (550 MHz cutoff) implemented in the open-source software Audacity (audacity.sourceforge.net).

Analyses

Behavioral analyses

Audio recordings acquired during fMRI scanning were processed to extract words and the language in which they were uttered. Despite our noise canceling procedures (see above), subjects' utterances remained quite noisy. Words were thus classified by an independent judge in three categories: (1) probably or definitely English, (2) could be English or German, (3) probably or definitely German. Words that could not be classified as either German or English (category 2) were excluded from further analysis. On average this excluded 8.9% of all trials. The percentage of items judged as sufficiently intelligible were comparable for English (46,6%) and German (44.5%).

Behavioral decoding of language choice. The sequence of languages across trials was analysed to make sure that the choice of languages was balanced, and that the active language in trial n could not be predicted on the basis of the language used in the preceding trial *n* - 1. This analysis was performed to assess whether the main multivariate pattern classification analysis on neural data may possibly rely on (informative) activation patterns generated during the execution phase of trial *n* - 1. The behavioral classification analysis is an estimate of the amount of information on trial *n* available from trial *n* - 1, and it constitutes a (generous) upper-bound of the potential role of this confound. Using cross validation, we tested the predictability of a language at time *t* given the language used at time *t* - 1. First, one run is left out and the probability of using language X at time *t* and language Y at time *t* - 1 is computed. Then the conditional probabilities of using language X at *t*, given that the language used at t - *1* is Y, was computed by:

$$P(X_t \vee Y_{t-1}) = \frac{P(X_t \cap Y_{t-1})}{P(Y_{t-1})}$$

From these probabilities a classifier was built. For example, if in a given trial English is used and if $P(G_t \mid E_{t-1})$ is larger than $P(E_t \mid E_{t-1})$ then the classifier will predict German for the next trial. This classifier was then used to predict the language in trials of the left out run. Using this calculation the accuracy of prediction in each left out run was tested.

Reaction Times Analysis. We computed reaction times between the onset of the line-drawing presentation and the utterance of the word. We evaluated possible differences between average RT for English and for German naming. Furthermore, we explored whether RTs were longer in those trials in which a different language was used with respect to the previous trial (switch trials) compared to those trials in which the language remained the same (stay trials).

Image Preprocessing

Preprocessing was performed using SPM12 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) and MATLAB 2013a. Images were slice-time corrected, re-aligned to correct for subject movements during acquisition, spatially normalized to the MNI space and smoothed with a Gaussian kernel of 6 mm FWHM. For multivariate analysis smoothing was not performed to preserve fine-grained patterns of individual voxels.

Neuroimaging analyses

Analyses were done using General Linear Model (GLM) and Multivariate Pattern Analysis (MVPA). As in our previous study on bilingualism (Reverberi et al., 2015) we analysed the effect of language and the effect of trial sequence during language choice, maintenance and execution phases. In each analysis we modeled data by using the canonical Hemodynamic Response Function (HRF). Furthermore, in order to have a more fine-grained temporal resolution we also implemented a Finite Impulse Response (FIR) in those analyses in which temporal information was important. In all analyses, beside the conditions of interest, we also included the movement parameters estimated during realignment as six covariates of no interest to remove possible artifacts due to head movement.

Language choice and maintenance analyses, HRF-based. In the HRF-based analyses four conditions were modeled at first-level using the Hemodynamic Response Function as implemented in SPM12: two conditions were used for English and German during the language choice and maintenance phase, and two conditions for English and German during the execution phase. Conditions for the language choice and maintenance phase were modeled by a boxcar function with onset corresponding to the beginning of the trial (green fixation cross), and duration corresponding to the time from the beginning of the trial to the onset of the execution phase (line-drawing presentation). Conditions for the execution phase were modeled as events with onset at the beginning of the execution phase. Catch trials were also included in this analysis. In this way, we could obtain 4 beta maps per subject per run (one map for each condition). These maps were submitted to a multivariate pattern analysis (MVPA) aimed at exploring which brain regions encoded the identity of the currently chosen language. MVPA was implemented using the searchlight approach (Kriegeskorte et al., 2006), allowing to systematically explore the local encoding of relevant information over the whole brain volume. The searchlight was a sphere with radius of 4 voxels. Linear Support Vector Machine (SVM) was chosen as classification algorithm. To prevent overfitting, a leave-one-out cross validation procedure was implemented so that SVMs were trained on data from all fMRI runs but one, and tested on the remaining one. Chance level was computed by Monte Carlo Simulation (Nichols Thomas E. & Holmes Andrew P., 2001). For generating the null distribution we permuted the class labels 100 times for each searchlight and for each subject. Permuted searchlights were then submitted to the same MVPA applied to non-permuted data. This step produced 1500 information maps over 15 participants. One chance accuracy map was randomly selected per every subject, and the selected 15 maps were averaged to one group accuracy map. Map selection and averaging was repeated 10000 times, resulting in a pool of 10000 chance accuracy maps providing an empirical chance distribution. The empirical chance distributions allowed the determination of a voxel-wise threshold which was used for a cluster search in the group chance accuracy maps (Stelzer, Chen, & Turner, 2013). To further check that the resulting significant clusters were informative, cluster confirmatory analyses was additionally applied (Etzel, Zacks, & Braver, 2013). Each cluster is tested for information as a region of interest by applying the same Monte Carlo Simulation procedure with 1000 permutations. If two separate clusters share a corner they are considered as one single ROI. Permutation-based MVPA analyses were performed by using PyMVPA package (Hanke et al., 2009).

Standard univariate analyses were also performed by using the same first-level model. We aimed at exploring possible activation differences between the two languages in each of the two task phases. We computed linear compound at the first level averaging activation related to German or English across all runs and in each task phase. These contrast images were then submitted to spatial normalization and spatial smoothing (FWHM = 6). The smoothed and normalised images were then submitted to a second-level analysis in SPM12 in order to test the presence of differences in activation between the two languages (two-tailed test).

Language choice and maintenance analysis, FIR-based. In the FIR-based analysis a Finite Impulse Response model was applied to the preprocessed images (Henson, 2003). Each condition was modeled using 16 time-bins of 2 seconds each. We used a relatively long FIR time series to remove from trial $n + 1$ (covered by the last time-bins) possible variance related to trial n. We considered two conditions corresponding to trials in which the chosen language was either German or English. Catch trials were excluded from this analysis. The onset time for FIR regressors was set at 4 s (2 time-bins) before the start of the trials to explore a possible presence of information about the language choice already at the end of the preceding trial. We assumed a hemodynamic delay of 4 s (two time-bins) from the onset of the cognitive event targeted and the hemodynamic response. Ten time-bins (from 1 to 10) were considered in the following analyses. These times bins covered three main time-windows: First, the pre-trial phase in which subjects are concluding the previous trial and have yet to choose which language to use in the next (FIR time-bins from 1 to 4, when the hemodynamic delay is considered); second, the language choice (time-bin 5) together with the maintenance phase (time-bins 6 to 8); third the execution phase (time-bins 9 and 10). An MVPA analysis identical to that described for the HRF-based analysis was run for the time-bins possibly related to language choice.

Trial sequence analysis. We investigated whether any brain region activates more in trials in which the chosen language was different from the preceding trial (switch trials) compared to trials in which the chosen language was the same as in the preceding trial (stay trials). We implemented HRF-based univariate analyses. The analyses performed were identical to those used in the language choice analysis except that the conditions corresponded to switch and stay trials instead of German vs. English.

In all whole-brain analyses we considered significant statistical tests having a voxel-level p-value $< .001$ (not corrected for multiple comparisons), and a cluster level p-value $< .05$ family-wise corrected for multiple comparisons (FWE correction; Friston, Holmes, Poline, Price, & Frith, 1996; Hayasaka, Phan, Liberzon, Worsley, & Nichols, 2004; Worsley et al., 1996). This restricted to a maximum of .05 the probability of falsely finding a cluster with a size equal or above a critical threshold.

Region of Interest (ROI) analyses. We ran ROI analyses to further assess the involvement of the brain regions found in the whole-brain analyses.

First, we wanted to assess whether brain regions that showed an effect in one task phase (e.g., language execution) would also show an effect in the other task phase (e.g. language choice and maintenance phase). Second, we wanted to formally compare the results of the present experiment with those of our previous experiment using forced language choice instead of free language decision. Besides language choice, other experimental details of the previous experiment were similar to the present experiment. Furthermore, the same individuals participated to the two experiments. Thus the formal comparison is across experimental sessions but within subjects.

The ROIs were based on the results of the whole-brain analyses. All voxels belonging to the same cluster as identified by SPM12 were assigned to the same ROI. Similar to the whole-brain multivariate analysis an SVM algorithm with leave-one-out cross validation was used for classification of the ROI data and the null distribution was estimated by 1000 permutation using Monte Carlo Simulation. In ROI analyses we considered significant statistical tests having a *p*-value < .05 corrected for multiple comparison (Bonferroni correction) when multiple ROIs were tested.

RESULTS

Language proficiency

All subjects reported German as their mother tongue and English as their strongest L2 in the LEAP-Q (see Tab. S1, supplementary online materials). The average reported proficiency in LEAP-Q for reading, listening and speaking in English was 8,3, 8.5, and 8.6 respectively (SD = 0.8, 0.9, and 0.6), while for German it was 9.9 (SD = 0.4) in all three areas of language use, on a scale from one to ten. Subjects' average exposure to English was 80%, or above. The average age of first exposure to English was 8.93 years (SD = 2.34) and 0.4 years (SD = 0.5) for German. Finally, since we were looking for subjects with an extensive vocabulary, included subjects could name correctly more than 90% of the pictures in the picture naming task, and showed a balanced choice between naming objects in English and German.

Behavioral results

Naming Reaction Times. Similar naming latencies ($t(14) = 0.04$; $p = 0.96$) were observed in German (1258 ms; SD = 198) and English (1262 ms; SD = 178). Switch trials led to similar latencies than stay trials (1269 ms (SD = 187) vs. 1257 ms (SD = 178), $t(14) = 0.17$; p = 0.86).

Behavioral decoding of language choice. Subjects chose to use English in 51% of all trials (SD = 4.1%). The frequency with which subjects chose to use English was not significantly different from their choice to use German (two-tailed t-test; $t(14) = 0.93$; $p = 0.37$). We evaluated whether it was possible to predict the language used at trial *n* based on the language used at trial *n* - 1. We trained a classifier to tell which language a subject would use based on the language used in the previous trial. The average accuracy of cross-validated prediction was 0.54 with SD = 0.05, which is significantly higher than chance level ($t(14)$ = 2.58, p = 0.01, one-tailed). This may reflect an individual's tendency to repeat a task more often than what would be expected by chance, even when the two tasks are chosen with equal frequency. This bias is frequently found in voluntary task switching paradigms (Arrington & Logan, 2004; Yeung, 2010). Importantly for our present research question, our analyses of the fMRI data (see below) indicated that language choice at trial n-1 did not affect the neural pattern of activity observed in a given trial n.

Neuroimaging analyses on chosen language

We explored whether subjects' choice to use German versus English elicited different average levels of activation in any region of the brain (univariate analyses), and whether local patterns of activation differed across languages (multivariate analyses). We considered two main task phases: in the first phase (language choice and maintenance phase, see Figure 1) subjects were instructed to choose and maintain the choice to use a particular language over a delay period. In the second phase (language execution phase), subjects named the presented object either in German (L1) or in English (L2).

Univariate analyses

Language choice and maintenance phase. While subjects were choosing which language to use in the upcoming trial, and maintaining this choice, we did not find significant activation differences between trials in which German and English were selected.

Language execution phase. During the actual picture naming phase, several brain regions were differentially engaged when subjects named the pictures in German as compared to English. Naming in German as compared to English activated more the angular gyrus bilaterally (Brodmann Area 39), the precuneus (BA 7) and left middle frontal gyrus (BA 46). On the other hand, the brain regions more active when using English as compared to German were the left inferior frontal gyrus (BA 47), left temporal pole (BA 21/22), parahippocampal gyrus (BA 35 and 37), pallidum, amygdala (BA 34), precuneus (BA 19) and cerebellum (Figure 2, Table 1).

Multivariate pattern analysis

We investigated whether and where local patterns of activation encoded the currently active language, both during the language decision and maintenance phase and during the language execution phase.

Language choice and maintenance phase. We first ran a multivariate pattern analysis (MVPA), applying a permutation procedure for empirically estimating the null distribution (Stelzer et al., 2013), and further checking for cluster informativeness with cluster confirmatory analyses (Etzel et al., 2013). We used beta images for maintenance/choice phase extracted from the HRF-model. We found that the local activity pattern in the medial prefrontal cortex (BA10) encoded information about the chosen language (Figure 3, Table 2). Notice that this region is more anterior to that

Next, we applied the same permutation procedure outlined above on beta images extracted from selected time-bins of the FIR model. We aimed at checking in which time-bin the information on the chosen language appeared first. We focused on the time-bins corresponding to language choice and maintenance phase (bins 5 to 8) and those immediately preceding choice (bins 3 to 4). Reliable information could be first found at time-bin 5. The time-bin 5 is roughly corresponding to the time of language choice (green fixation). Information in bin 5 was again encoded in medial prefrontal cortex (center at [-15 54 17], BA 10). The availability of information already in bin 5, besides showing its involvement in language choice, indirectly confirms that participants complied with the instructions of choosing language in the task decision phase.

Finally, we tested whether the performance of the classifiers relying on the fMRI activation patterns in trial n also exploited the information associated with the identity of the chosen language in trial $n - 1$. If this were the case one should expect a positive between-subject correlation between the performance of the fMRI classifier and that of the classifier based on the decision in trial $n - 1$ (see section "behavioral decoding of language choice"). We checked the medial prefrontal cluster in BA10 (Figure 3, Table 2). The across-subject correlation between decoding accuracy in BA10 and behavioural decoding accuracy was not significantly different from 0 (pearson correlation: r = 0.0345; $p = 0.75$).

Language execution phase. Multivariate pattern analysis was also run during the language execution phase, using the beta images extracted from the HRF-model. During language execution the chosen language could be decoded from local activation patterns in the left inferior frontal gyrus (mainly BA 44/45/47), the right supramarginal gyrus (BA 40), left inferior parietal lobule

(BA 40), and left temporal gyrus (BA 37/21) (Figure 4, Table 3). Again, in none of these clusters the performance of the classification based on neural data was significantly correlated with the performance based on behavioral data (cluster at BA 45/46 : r = -0.1008, *p* = 0.34, cluster at BA 40/48 r = 0.0618, *p* = 0.56 cluster at BA 21/22/37 r = 0.1454, *p* = 0.17, cluster at BA 44/48 r = 0.0776, *p* = 0.47, cluster at r = 0.0007, *p* = 0.99, and cluster at BA40 r = -0.2, *p* = 0.059).

ROI analyses. We tested whether any of the relevant regions found either during language choice and maintenance, or during execution are also involved during the other task phase. In the whole-brain univariate analysis, we only found differences between German and English during the execution phase (Tab. 1, Fig. 2). 13 ROIs were defined based on these results. We then tested whether in any of these regions (defined on execution phase) the average activity in the ROIs would differ between German and English during language choice and maintenance. After correcting for multiple comparisons, we did not find a significant difference in any of the ROIs.

By using MVPA we found one informative cluster during language choice and maintenance phase (Tab. 2, Fig. 3), and five during execution phase (Tab. 3, Fig. 4). Thus, six ROIs overall were defined. We then extracted the average accuracy level from these ROIs. We found that in the ROIs derived from the execution phase (Tab. 3) the classifier performance was not different from chance in the language choice and maintenance phase (all *p*s > 0.05 uncorrected). We repeated the same procedure for the only ROI derived from the choice/maintenance phase. We found that the classifier performance was not different from chance in the language execution phase (*p* = 0.26). Thus, overall, the brain regions identified by MVPA both in choice and execution are task phase specific.

Neuroimaging analyses on effects of trial sequence

We tested whether trials in which subjects used a different language with respect to the previous trial (switch trials) show higher activation in any part of the brain compared to those trials in which subject continued to use the same language (stay trials). The analysis was performed both in the language choice and maintenance phase and in the language execution phase (HRF model). No significant differences were found.

Comparison with forced-choice experiment

We tested whether the brain regions encoding the chosen language when the subjects are free to decide the language to use (current experiment), also encode the active language when the choice is forced by the experimenter (Reverberi et al., 2015). To this aim we used the ROI of the medial prefrontal regions encoding language in the current experiment and we applied it to our previous data-set. Notice that the previously reported experiment (Reverberi et al., 2015) applied a very similar experimental paradigm to the one implemented here, the only major difference being the presence of free vs forced choice of the language. Furthermore, the subjects involved in the two experiments were the same. By contrast with the present experiment, we found that the medial prefrontal cortex did not encode the active language when the choice is not free ($p = 0.834$). The active language was rather encoded in the occipital cortex (BA 18/19), premotor cortex (BA 6) and caudate nucleus (see supplementary online material for details).

DISCUSSION

In communicative contexts in which the two languages can be used at will without compromising communication, choosing a certain language can be described as a self-determined action made by speakers themselves. Once this choice has been made, bilinguals have to act upon their choice and take care of potential interference from the non-selected language. Yet, in most experimental studies on bilingual language production speakers do not freely choose which language to use. Moreover, most studies typically do not differentiate between two distinct phases, i.e., the phase during which a language is chosen and the phase during which verbal output is produced (execution). In the present study we tackled both shortcomings by (1) allowing bilinguals to freely choose which language to use in each single naming event, and (2) by temporally separating the phase in which bilinguals need to choose and maintain their choice to speak a particular language from the phase in which they need to speak in the selected language. Through the use of Multivariate Pattern Analysis (MVPA) we were able to identify the brain regions encoding a speaker's language choice, independently of the actual execution of such choice.

We found that language choice and language execution involve different brain regions. While language choice relies on a region related to free choice also in other non-linguistic domains, language execution relies on the so-called language control network, a network of brain regions thought to ensure that a language-specific lexical item is selected in spite of potential competition between languages.

Language choice and maintenance

During this task phase participants freely decide which language to use for naming the upcoming picture and then keep active this decision until the actual picture is shown. Univariate analyses comparing the activation elicited when the choice was L1 (German) vs. L2 (English), failed to detect any difference between them. In contrast, multivariate pattern analyses found that the chosen language was encoded predominantly in the medial prefrontal cortex (BA10). Information on the chosen language was available in this area as early as the presentation of the language neutral cue choice.

Interestingly, studies on other types of free choices have reported overlapping regions in medial prefrontal cortex, hence underlining its role in domain-general executive control. Decisions in these studies ranged from simple alternative motor actions to complex rule sets, choices between alternative arithmetical operations, different consumer products, and strategies (Hampton & O'Doherty J, 2007; Haynes et al., 2007; Schuck et al., 2015; Soon, Brass, Heinze, & Haynes, 2008; Soon, He, Bode, & Haynes, 2013; Tusche, Bode, & Haynes, 2010; Wisniewski, Reverberi, Tusche, & Haynes, 2015; Zhang, Kriegeskorte, Carlin, & Rowe, 2013). For example, Wisniewski and collaborators (Wisniewski et al., 2015) asked participants to freely choose, on a trial by trial basis, one out of three rule sets (where a rule is for example "if you see a musical instrument then press the red button"), while the difficulty associated with each rule set changed dynamically. Wisniewski and collaborators found that the medial prefrontal cortex encoded which rule set a participant decided to apply in the succeeding trial. The present study is consistent with this previous evidence, and further corroborates the idea that the medial prefrontal cortex is involved in processing free decisions in a wide range of domains. Of course, the possibility remains that within the medial prefrontal cortex there are distinguishable neural populations mediating language choice specifically. To investigate this possibility, an additional, nonlinguistic control condition would be necessary. However, the convergence in the literature across domains suggests that language choice in bilinguals does not involve processes that are language specific but rather domain general.

Other studies, using univariate analyses, have reported higher activation in the medial prefrontal cortex in free-choice tasks compared to forced-choice tasks (Brass & Haggard, 2008; Brass, Lynn, Demanet, & Rigoni, 2013; Demanet, De Baene, Arrington, & Brass, 2013; Forstmann, Brass, Koch, & von Cramon, 2006; Orr & Banich, 2014; Walton, Devlin, & Rushworth, 2004; Zhang et al., 2015). Most relevant in the context of language production is the paper by Zhang and collaborators (Zhang et al., 2015). In this study bilingual participants were asked to name digits either in Chinese or in English. In one experimental condition participants were free to choose which language to use, while in the other condition participants were cued to use a particular language. A large fronto-parietal network, including the anterior cingulate and the SMA, was more active during free language selection than during forced language selection. Compared to those of Zhang and collaborators, our experimental paradigm and findings differ along several dimensions.

From the cognitive point of view, our experimental paradigm allowed to tease apart language choice from actual language execution, avoiding collapsing two task phases likely involving different neural structures (Reverberi et al., 2015). Furthermore, the use of multivariate analyses allowed us to identify brain regions specifically encoding the selected language. We found that only one brain region (medial PFC) is critical during the choice phase. This region lies more anterior and inferior compared to the medial prefrontal region active in Zhang et. al's study. In analogy with evidence on free decisions in the motor domain (Soon et al., 2008), we may speculate that our paradigm and analysis strategy identified signals from an earlier cognitive stage than that observed by Zhang and collaborators. Alternatively, we may hypothesize that the network observed by Zhang and collaborators is recruited only when free language choice interacts with further executive requirements related to the immediate use of the chosen language.

Language execution

After choosing the language in which the verbalization will take place and maintaining such choice during 6 seconds, the actual target picture was presented and participants entered into the execution phase. During this language execution phase, a large network of brain regions was found to be involved. Multivariate analyses testing whether local patterns of activation differed across languages showed the involvement of the left Broca's area (BA 44/45), the right supramarginal gyrus and the right inferior temporal gyrus. Univariate analyses, contrasting activation associated with speaking one language to the other, identified multiple regions displaying differential activation between the two languages, including the left inferior frontal lobe (BA 44 and 47), the basal ganglia, and the angular and inferior parietal gyrus bilaterally (BA 40/39). Importantly, none of the regions involved in the language execution phase encoded the selected language during the choice and maintenance phase, showing that language choice and language execution rely on distinct neural and cognitive processes.

The brain regions involved in language execution in this study largely overlap with those reported in our previous study (Reverberi et al., 2015). We argued that these areas relate to the cognitive control processes that regulate the use of one language over the other. Beyond showing the robustness of our previous observations, the present experiment further shows that the brain network supporting language execution remains the same irrespectively of whether the language was cued by the experimental setting or freely chosen by the speaker. The independence of the language execution network from the way the active language has been selected agrees with its hypothesized cognitive role, namely, being in charge of resolving competition between languages when language-specific lexical items need to be retrieved (Abutalebi & Green, 2007).

Limits

The sample size of the experiment is limited (n = 15). Unfortunately, some of the subjects who participated in our forced-choice experiment (Reverberi et al., 2015) were no longer available for a second experimental session. Notwithstanding the somewhat smaller sample size, we preferred not to involve new subjects to keep the homogeneity between the two experimental samples. Given the limited sample, we are cautious in interpreting whole-brain negative results, if they are not confirmed by a more sensitive ROI analysis. In the case of ROI analyses, with $\alpha < .05$ (no multiple comparisons correction) we have 80% power for effect size d = .68. When a more liberal $\alpha < .1$ is used, we would have 80% power with an effect size d = .56. Besides, it is reassuring that our findings for the language execution phase are consistent with those reported in our previous study with a larger sample size (Reverberi et al., 2015), while results on language choice are consistent with previous studies on free decision. Another limitation refers to the fact that we only tested late bilinguals and consequently we cannot make broad generalizations regarding whether the observed results would be replicable in other bilingual speakers, especially on those that have acquired the language early in life.

Conclusions

We explored the neural structures involved in language choice when bilinguals are free to choose the target language, and how these neural structures relate to the language control network active during language use. The medial prefrontal cortex encoded the chosen language prior to speaking. By contrast, during language execution, a larger brain control network was recruited including the left inferior frontal lobe, the basal ganglia, and the angular and inferior parietal gyrus bilaterally. None of the language control regions were involved during language choice. Thus, the neural and cognitive processes related to bilingual language choice are different from those related to language control during language execution. Furthermore, our findings support both the idea that medial prefrontal cortex is a domain general region critical for free choice, and that bilingual language choice relies on domain general processes.

ACKNOWLEDGMENTS

CR and SSA were supported by the PRIN grant 2010RP5RNM_001 from the Italian Ministry of University; AC was supported by two grants from the Spanish Government, PSI2011-23033, PSI2014-52181-P, a grant from the Catalan government (AGAUR SGR 268), and a grant from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Cooperation grant agreement nº 613465 - AThEME).

REFERENCES

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242–275. doi:10.1016/j.jneuroling.2006.10.003

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition*, *19*(4), 689–698. doi:10.1017/S1366728916000225

Arrington, C. M., & Logan, G. D. (2004). The cost of a voluntary task switch. *Psychol Sci*, *15*(9), 610–615. doi:10.1111/j.0956-7976.2004.00728.x PSCI728 [pii]

Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *14*(4), 319–325. doi:10.1177/1073858408317417

Brass, M., Lynn, M. T., Demanet, J., & Rigoni, D. (2013). Imaging volition: what the brain can tell us about the will. *Experimental Brain Research*, *229*(3), 301–312. doi:10.1007/s00221-013-3472-x

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical Selection in Bilinguals: Do Words in the Bilingual's Two Lexicons Compete for Selection? *Journal of Memory and Language*, *41*(3), 365–397. doi:10.1006/jmla.1999.2651

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. doi:10.1016/j.jml.2004.02.002

Demanet, J., De Baene, W., Arrington, C. M., & Brass, M. (2013). Biasing free choices: The role of the rostral cingulate zone in intentional control. *NeuroImage*, *72*, 207–213. doi:10.1016/j.neuroimage.2013.01.052

Etzel, J. A., Zacks, J. M., & Braver, T. S. (2013). Searchlight analysis: Promise, pitfalls, and potential. *NeuroImage*, *78C*, 261–269. doi:10.1016/j.neuroimage.2013.03.041

Forstmann, B. U., Brass, M., Koch, I., & von Cramon, D. Y. (2006). Voluntary selection of task sets revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, *18*(3), 388–398.

Friston, K. J., Holmes, A., Poline, J. B., Price, C. J., & Frith, C. D. (1996). Detecting activations in PET and fMRI: levels of inference and power. *NeuroImage*, *4*(3 Pt 1), 223–35. doi:10.1006/nimg.1996.0074

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 640–665. doi:10.1037/a0014981

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(02), 67–81. doi:10.1017/S1366728998000133

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. doi:10.1080/20445911.2013.796377

Hampton, A. N., & O'Doherty J, P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(4), 1377–1382. doi:0606297104 [pii] 10.1073/pnas.0606297104

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Olivetti, E., Fründ, I., Rieger, J. W., ⋯ Pollmann, S. (2009). PyMVPA: A Unifying Approach to the Analysis of Neuroscientific Data. *Frontiers in Neuroinformatics*, *3*, 3.

Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *NeuroImage*, *22*(2), 676–87. doi:10.1016/j.neuroimage.2004.01.041

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Review Neuroscience*, *7*(7), 523–534.

Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Current Biology*, *17*(4), 323–328.

Henson, R. N. A. (2003). Analysis of fMRI time series. In R. S. J. Frackowiak, K. J. Friston, C. Frith, R. Dolan, K. J. Friston, C. J. Price, ⋯ W. D. Penny (Eds.), *Human Brain Function* (2nd ed.). Academic Press.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–685. doi:nn1444 [pii] 10.1038/nn1444

Kleinman, D., & Gollan, T. H. (2016). Speaking Two Languages for the Price of One Bypassing Language Control Mechanisms via Accessibility-Driven Switches. *Psychological Science*, *27*(5), 700–714. doi:10.1177/0956797616634633

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3863–3868. doi:0600244103 [pii] 10.1073/pnas.0600244103

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, *9*(02), 119–135. doi:10.1017/S1366728906002483

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*, *50*(4), 940–967. doi:10.1044/1092-4388(2007/067)

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language*, *40*(1), 25–40. doi:10.1006/jmla.1998.2602

Nichols Thomas E., & Holmes Andrew P. (2001). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping*, *15*(1), 1–25. doi:10.1002/hbm.1058

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430. doi:S1364-6613(06)00184-7 [pii] 10.1016/j.tics.2006.07.005

Orr, J. M., & Banich, M. T. (2014). The neural mechanisms underlying internally and externally guided task selection. *NeuroImage*, *84*, 191–205. doi:10.1016/j.neuroimage.2013.08.047

Reverberi, C., Kuhlen, A., Abutalebi, J., Greulich, R. S., Costa, A., Seyed-Allaei, S., & Haynes, J.-D. (2015). Language control in bilinguals: Intention to speak vs. execution of speech. *Brain and Language*, *144*, 1–9. doi:10.1016/j.bandl.2015.03.004

Schuck, N. W., Gaschler, R., Wenke, D., Heinzle, J., Frensch, P. A., Haynes, J.-D., & Reverberi, C. (2015). Medial Prefrontal Cortex Predicts Internally Driven Strategy Shifts. *Neuron*, *86*(1), 331–340. doi:10.1016/j.neuron.2015.03.015

Snodgrass, J. G., & Vanderwart, M. (1980). Standardized set of 260 pictures: norms for name ageement, image agreement, familiarity and visual complexity, *6*, 174–215.

Soon, C. S., Brass, M., Heinze, H. J., & Haynes, J. D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*(5), 543–545. doi:nn.2112 [pii] 10.1038/nn.2112

Soon, C. S., He, A. H., Bode, S., & Haynes, J.-D. (2013). Predicting free choices for abstract intentions. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1212218110

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. *NeuroImage*, *65*, 69–82. doi:10.1016/j.neuroimage.2012.09.063

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., ⋯ Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250. doi:10.1016/j.jml.2004.03.002

Tusche, A., Bode, S., & Haynes, J.-D. (2010). Neural responses to unattended products predict later consumer choices. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(23), 8024–31. doi:10.1523/JNEUROSCI.0064-10.2010

Walton, M. E., Devlin, J. T., & Rushworth, M. F. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, *7*(11), 1259–1265. doi:nn1339 [pii] 10.1038/nn1339

Wisniewski, D., Reverberi, C., Tusche, A., & Haynes, J.-D. (2015). The Neural Representation of Voluntary Task-Set Selection in Dynamic Environments. *Cerebral Cortex*, *25*(12), 4715–4726. doi:10.1093/cercor/bhu155

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*, *4*(1), 58–73.

Yeung, N. (2010). Bottom-up influences on voluntary task switching: the elusive homunculus escapes. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *36*(2), 348–362. doi:10.1037/a0017894

Zhang, J., Kriegeskorte, N., Carlin, J. D., & Rowe, J. B. (2013). Choosing the Rules: Distinct and Overlapping Frontoparietal Representations of Task Rules for Perceptual Decisions. *The Journal of Neuroscience*, *33*(29), 11852–11862. doi:10.1523/JNEUROSCI.5193-12.2013

Zhang, Y., Wang, T., Huang, P., Li, D., Qiu, J., Shen, T., & Xie, P. (2015). Free Language Selection in the Bilingual Brain: An Event-Related fMRI Study. *Scientific Reports*, *5*, 11704. doi:10.1038/srep11704

Figure 1 Timeline of one trial of the experimental task. Participants were first required to choose the language to use in the current trial (choice phase, lasting 2 s). Then they had to maintain this choice (maintenance phase, lasting 1 or 6 s), and finally use the chosen language to name an object. In each trial, two main phases were considered for the analysis: the language choice/maintenance phase and the language execution phase. In the former, the subjects had to choose and maintain the relevant language, while in the latter they had to use the chosen language to name the presented object.

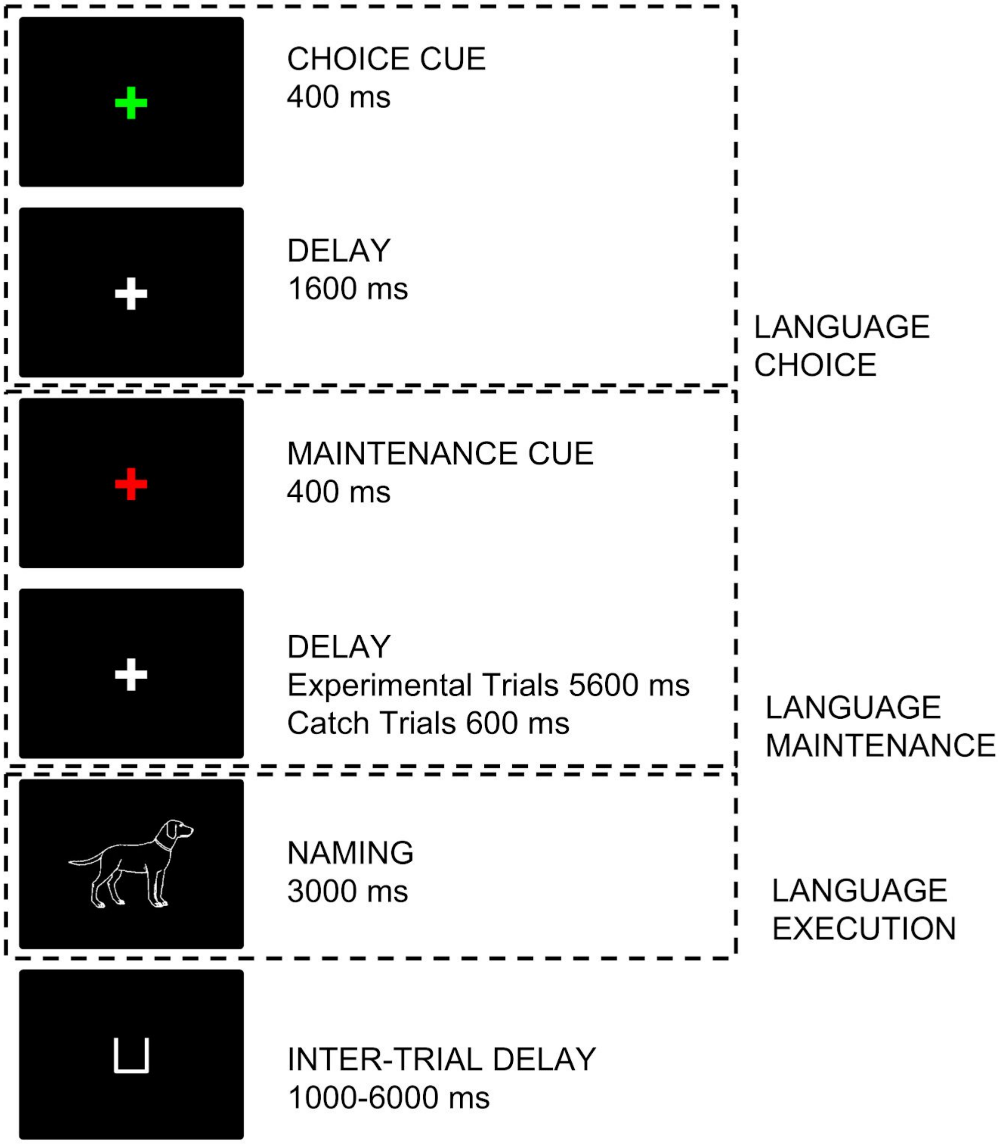

Figure 2 Univariate analysis: Activation differences between languages during language execution. The regions that are more activated for English during execution are shown in red and those that are more activated for German are shown in yellow.

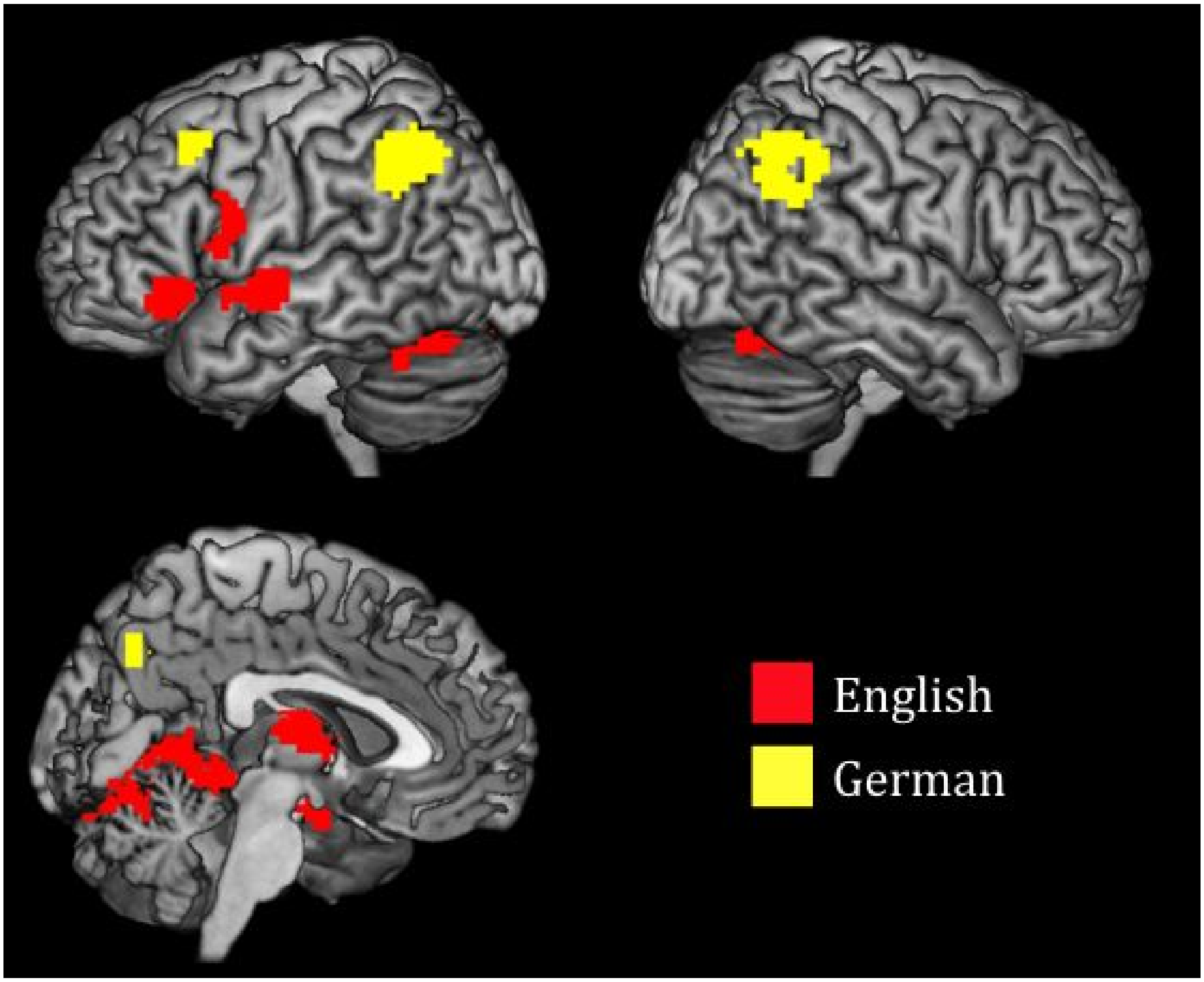

Figure 3 Multivariate Pattern Analysis: Brain regions encoding the chosen language during the language choice and maintenance phase.

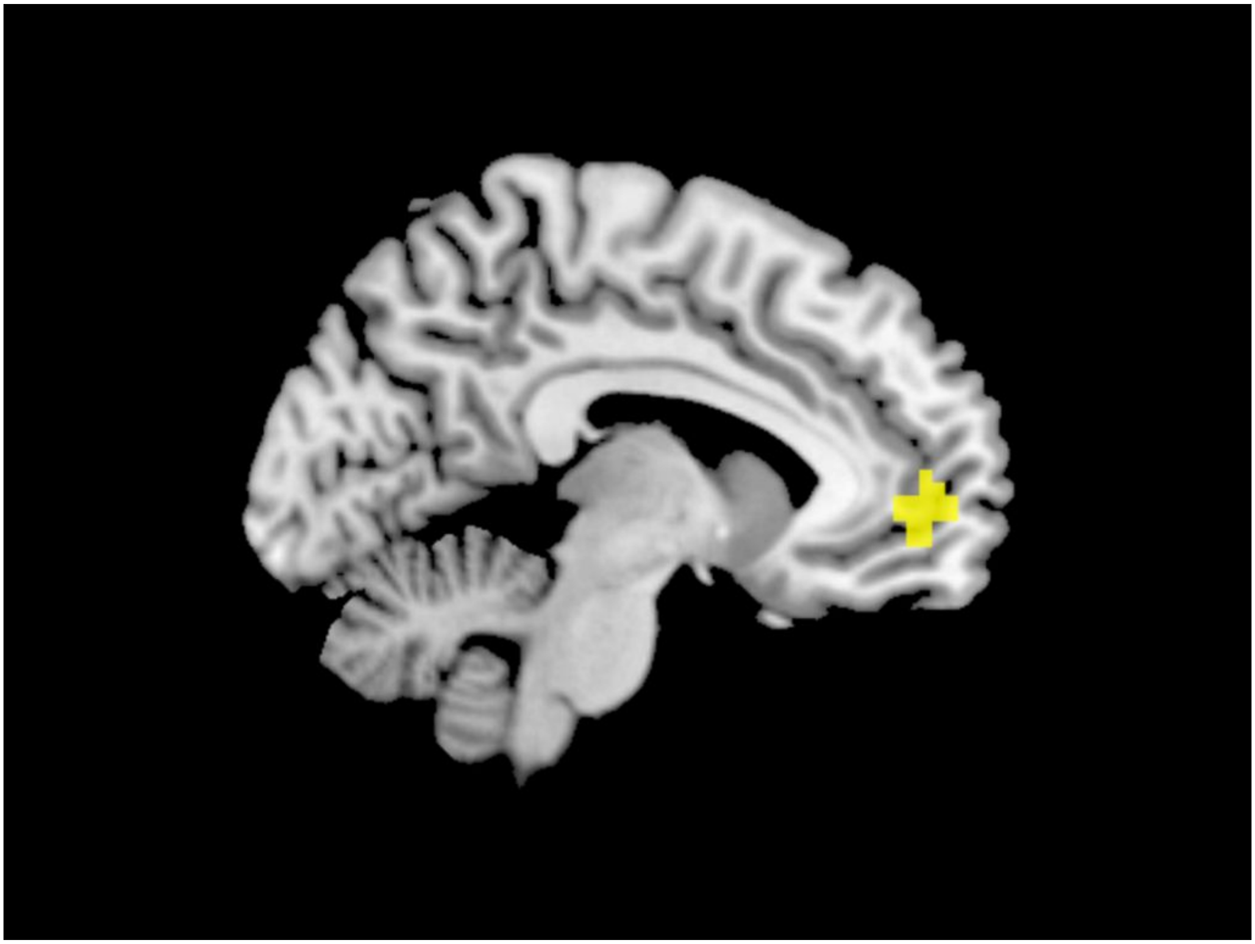

Figure 4 Multivariate Pattern Analysis: Brain regions encoding the chosen language during the language execution phase.