Language control in bilinguals: Intention to speak vs. execution of speech

Carlo Reverberi [a,b,*,1], Anna Kuhlen [c,d,e,1], Jubin Abutalebi [f], R. Stefan Greulich [d,e], Albert Costa [g,h], Shima Seyed-Allaei [a,b], John-Dylan Haynes [c,d,e]

[a] Department of Psychology, Università Milano – Bicocca, Milan, Italy
[b] NeuroMI – Milan Center for Neuroscience, Milan, Italy
[c] Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany
[d] Bernstein Center for Computational Neuroscience Berlin and Charité, Berlin, Germany
[e] Berlin Center of Advanced Neuroimaging, Charité – Universitätsmedizin Berlin, Berlin, Germany
[f] Department of Clinical Neurosciences, University San Raffaele and Scientific Institute San Raffaele, Milan, Italy
[g] Center of Brain and Cognition, CBC, Universitat Pompeu Fabra, Spain
[h] Institució Catalana de Recerca i Estudis Avançats (ICREA), Spain

ARTICLE INFO

ABSTRACT

Bilinguals require a high degree of cognitive control to select the language intended for speaking and inhibit the unintended. Previous neuroimaging studies have not teased apart brain regions for generating the intention to use a given language, and those for speaking in that language. Separating these two phases can clarify at what stage competition between languages occurs. In this fMRI study German–English bilinguals were first cued to use German or English. After a delay, they named a picture in the cued language. During the intention phase, the precuneus, right superior lateral parietal lobule, and middle temporal gyrus were more activated when participants had to update the currently active language. During language execution activation was higher for English compared to German in brain areas associated with cognitive control, most notably the anterior cingulate and the caudate. Our results suggest two different systems enabling cognitive control during bilingual language production.

1. Introduction

Speaking requires a high degree of cognitive control in order to monitor and correct for possible errors. In the case of bilingual speakers, the demand for cognitive control is even greater since they must also monitor and correctly select the intended language for discourse. Indeed, before bilinguals speak, they need to decide which language to use. Often, cues from the conversational context indicate which language is appropriate (Grosjean, 1982): for example, in a bilingual work environment it may be customary to hold formal presentations in English, while informal conversations may typically take place in German. Other cues that inform the intention to speak a particular language may be who the conversational partner is, or which language was just used (e.g., Schwartz & Kroll, 2006). In this study we will address how

cognitive control processes assist the bilingual brain in encoding and acting upon the intention to use one of two languages.

Models of bilingual speech production generally assume that bilinguals have at least two lexical representations attached to the same concept (De Groot, 1993; Francis, 1999; Kroll & Stewart, 1994). Thus, in order to name an object in the language appropriate for a given context, bilinguals must select the intended language and avoid interference from the unintended language. Cognitive control mechanisms are assumed to manage and resolve potential competition among distinct language systems (e.g., Abutalebi & Green, 2007; Costa & Santesteban, 2004; Green, 1986, 1998; Luk, Green, Abutalebi, & Grady, 2012; Price, Green, & von Studnitz, 1999). Neuroimaging studies have been successful at identifying a set of cortical and subcortical regions responsible for selecting one language over the other (Abutalebi & Green, 2007; Crinion et al., 2006; Garbin et al., 2011; Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001; Luk et al., 2012; Rodriguez-Fornells et al., 2005). Most notably, a bilingual language control network has been associated with the pre-Supplementary Motor Area/Anterior Cingulate Cortex (pre-SMA/ACC), prefrontal

cortex and the left caudate (Abutalebi & Green, 2007; Abutalebi et al., 2013).

While there is general agreement that control mechanisms promote the activation of the intended and prevent interference from the unintended language, the time course of when exactly control mechanisms are exerted upon the unintended language is disputed. In most occasions the decision on which language to speak is set at the preverbal stage, that is at the stage at which pragmatic cues about the communicative context become available, but before the specific content of the communication is set. Accordingly, one psycholinguistic account proposes that the (preverbal) intention to speak one language rather than the other is sufficient to unambiguously select words in the intended language during the verbal phase (La Heij, 2005). Following a similar account, contextual cues are proposed to bias production toward the intended language to the extent that only words in the target language are considered for selection (Costa & Caramazza, 1999; Costa, Miozzo, & Caramazza, 1999; Finkbeiner, Gollan, & Caramazza, 2006; Roelofs, 1998). According to this proposal, competition between different lexical choices only occurs within the target language. On the other hand, a different group of cognitive models assume that lexical representations within the intended language compete actively with alternatives in the unintended language until a rather late stage (e.g., Green, 1986, 1998; Kroll, Bobb, & Wodniecka, 2006; Poulisse & Bongaerts, 1994). Eventually, these alternatives are inhibited (e.g., Green, 1998). According to this proposal, the intention to speak one language is not sufficient to activate selectively the lexical representations in the intended language only.

Previous imaging studies have not fully disentangled neural activity associated with the process of forming an intention to use a particular language from neural activity associated with the process of acting upon this intention (i.e., speaking). Typically, the experimental paradigms used in previous fMRI studies presented the cue to use a particular language, e.g., English, in close temporal proximity to the prompt to speak, e.g., the presentation of an object to be named (Abutalebi et al., 2008, 2013). Given the limited temporal resolution of fMRI, such paradigms do not allow a separate investigation of the brain regions involved in processing the cue (intending to use a given language) and those involved in actually retrieving the lexical representation. Importantly, while the processing of the language cue is preverbal (speakers do not know yet *what* they will say, but only which language to use), naming an object is obviously verbal. Hence, if the instruction about which language to use coincides with the instruction on the linguistic content, it is impossible to investigate separately the stage at which speakers form the preverbal intention to use a particular language and the verbal stage at which speakers seek lexical access. An explicit separation of the processes involved in these two task phases in a group of bilinguals would clarify at what stage competition between languages occurs during language production. Furthermore, it would also allow to investigate whether cognitive control processes are involved at both stages, and, if so, whether they act via different brain networks.

While typically speakers speak directly after forming the intention to speak, task cueing paradigms ask participants to represent and maintain active the intention to perform an upcoming task over a delay period (e.g., Brass & von Cramon, 2002; Dosenbach et al., 2006; Haynes et al., 2007; Reverberi, Görgen, & Haynes, 2012a,b; Sakai & Passingham, 2003). Neural activity measured during this delay period is thought to reflect cognitive preparation for performing the upcoming task. Task cueing paradigms have been successfully applied to investigate the neural processes underlying the intention to speak in monolinguals. For instance, in anticipation of linguistic material for articulation, subjects activate the entire language production network 2–4 s prior to speaking,

including the frontopolar (BA 10) and anterior cingulate cortices, the supplementary motor areas (SMAs), and the caudate nuclei (Gehrig, Wibral, Arnold, & Kell, 2012; Kell, Morillon, Kouneiher, & Giraud, 2011). In the case of bilingual language production, it remains unknown whether preparatory processes associated with speech production include representations specific to the language the speaker chooses to speak.

In the present functional magnetic resonance imaging (fMRI) study we investigate how the bilingual brain encodes and acts upon the intention to speak one of two languages. We apply a task-cueing paradigm to separate neural activation associated with forming the intention to speak from neural activation associated with actually retrieving specific lexical items in the selected language. During scanning, German–English bilingual participants were first cued to use a particular language (German or English). They were then asked to represent and maintain the intention to use that language over a delay period (*language intention phase*). After this delay, participants were presented a picture of an everyday object, which they were asked to name in the cued language (*language execution phase*). This separation of language intention and language execution may potentially allow us, as aforementioned, to address whether control processes necessary to resolve competition between the two languages are already active at the preverbal stage, or, only at the verbal stage, i.e., at the time of lexical access. Hence we compared neural activation associated with German to English single word production during the intention and the execution phase. If having the intention to speak a particular language triggers the suppression of the unintended language (i.e., favors representation associated with the intended language) we should be able to detect systematic activation differences in brain areas associated with language control between trials in which speakers intend to speak English (less dominant language) and trials in which speakers intend to speak German (more dominant language). However, if both languages compete actively until a rather late stage in the production, we should not be able to detect difference between German and English trials during the intention phase, but instead only during the execution phase. Furthermore, we studied the brain structures involved in the updating of the active language by comparing trials in which the cued language changed from the preceding trial (switch trials) to those in which language remained the same (stay trials). We expect that switch trials will result in a pronounced pattern of activation related to cognitive control as these trials require the speaker to re-select the language that will be produced.

2. Methods

2.1. Participants

German–English bilinguals were recruited with flyers and emails. The advertisement asked for German native speakers with a very good knowledge of English. Volunteers who reported that English was their second mother tongue language (early bilinguals) or who reported high fluency in a third language were excluded. The remaining volunteers were invited for a first experimental session, which included an extended assessment of their language proficiency. The final subject group ($n = 23$) was selected on the basis of the language proficiency tests. They were all high proficiency late bilinguals (see Section 2.2 for details). Two subjects were further excluded because they asked to abort scanning. The remaining 21 subjects with full scanning data (age 23.1 ± 3.1, 12 female) were right handed, as assessed by the Edinburgh handedness inventory (mean = 67.7, SD = 9.8 with 44 as the lowest reported score, Oldfield, 1971). All subjects had normal or corrected to normal vision. No subject reported a neurological or

psychiatric history, or a language or learning impairment. The experiment was approved by the ethics committee of the department of psychology of the Humboldt University. Subjects received monetary reward for their participation.

2.2. Language assessment and selection criteria

Language proficiency was assessed by means of a German version of the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian, Blumenfeld, & Kaushanskaya, 2007), and a picture naming task using a procedure similar to the experiment itself (see below). We selected subjects reporting German as their mother tongue and English as their strongest L2 in the LEAP-Q. Furthermore, subjects with an average exposure to English language below 20% were excluded from the study. We also excluded subjects reporting a third language with a daily usage over 10% to avoid possible interference from a third strong language. Finally, since we were looking for subjects with an extensive vocabulary, we excluded subjects who correctly named less than 90% of the pictures in the picture naming task.

2.3. Stimuli

During the main experiment subjects were required to name 120 black and white line drawings of objects. These stimuli were 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. Of all indices available per each object we considered naming reaction times and word agreement (i.e., the proportion of subjects using a specific word to name an object). 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 drawings associated with words with different number of syllables in the two languages. From the remaining drawings we selected for the experiment the 120 items allowing to maximize the word agreement and, secondarily, to minimize the reaction times. Frequency of the target words was highly similar in the two languages: mean word frequency in German was 2.78 (SD = 1.57), while in English it was 2.64 (SD = 1.52).

Four abstract visual cues were used to code the active language. Two cues were associated with each language. The cue–language association was balanced over subjects.

2.4. Experimental procedure

The fMRI experiment was divided into six runs, each comprising of 40 trials. Each fMRI run lasted on average 9.5 min. During each experimental trial (Fig. 1), subjects were required to name a line drawing of an object using the currently active language. Each trial began with the presentation of a fixation cross for 500 ms. The fixation cross was followed by an abstract visual cue instructing the subject on which language to use for naming the upcoming picture. Cue presentation lasted 1 s, and it was followed by blank screen with a fixation cross lasting 8 s. During this delay phase subjects had to maintain the active language and prepare to respond in the target phase. Eight catch trials with a shorter delay (2, 4, or 6 s) were randomly introduced into each run, to ensure that subjects were keeping active the relevant language and they were preparing for task execution for the whole delay. During the target phase, a line-drawing was displayed for 3 s, within which subjects had to generate a response. The verbal response was recorded by an MRI compatible microphone. The 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 target
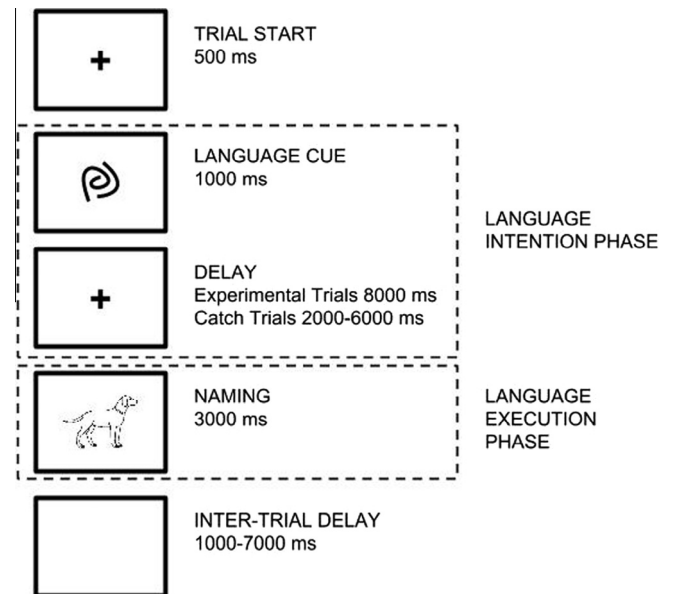
Fig. 1. Timeline of one trial of the experimental task. Subjects were required to name an object using the language active in the trial. The active language was conveyed by a visual cue presented at the beginning of the trial. In each trial two main phases were considered: the language intention phase and the language execution phase. In the former the subjects had to retrieve and represent the active language, while in the latter they had to use the active language to name the presented object.

phase, a blank screen was presented for a variable duration ranging from 1 to 7 s (average 3.5 s, exponential distribution with mode 1). Matlab and the Cogent toolbox (http://www.vislab.ucl.ac.uk/cogent.php) were used to present the stimuli.

The sequence of languages across trials was pseudo-randomized so that the same language was never active more than three times in a row. Trials in which the cued language changed from the preceding trial were called "switch trials" while those in which the language remained the same are called "stay trials". The generated trial sequences were further checked to make sure that the active language in trial n could not be predicted on the basis of other features of the trial n or the immediately preceding trial $n - 1$.

Subjects underwent further behavioral training before fMRI scanning. The training involved two major steps. The first step was aimed at making sure that subjects learned the association between the visual cue and the language to be used. In each trial subjects were presented with one of the four visual cues and had to select the language associated to that cue. The training continued until the subject correctly identified the language associated with every cue for at least eight times in a row. The time available for identifying a cue was limited and decreased gradually over the course of the training procedure from 2 to 1 s. This training phase had a variable duration depending on subject performance. All subjects reached the criterion within 10 min. In the second step of the training phase subjects could familiarize with a task identical to the one used in fMRI in all respect but for the target stimuli used. Ten line-drawing pictures were administered to all subjects. This training phase lasted about 5 min. Besides training with the experimental task, the subjects could also practice uttering answers in a microphone with minimal jaw movements.

2.5. 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, 64×64; and a field of view (FOV), 192×192 mm; thus yielding an in-plane voxel resolution of 3 mm^2, resulting in a voxel size of $3 \times 3 \times 3.75$ mm. A T1-weighted structural dataset was also collected. The parameters were as follows: TR, 2000 ms; TE, 2.52 ms; matrix size, 256×256; 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 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. The noise cancellation system allowed recordings of human voice with a quality good enough for understanding the words produced. Because of a technical problem, however, the noise cancellation system did not work for 11 subjects. All audio recordings were further cleaned in a post-acquisition phase with a noise cancellation algorithm and a low pass filter (550 MHz cut-off) implemented in the open-source software Audacity (audacity.sourceforge.net). The suboptimal quality of the recording in some of the subjects prevented the identification of the words uttered in the scanner, but it still allowed the detection of the beginning of the utterance (relevant for reaction time evaluation).

2.6. Analyses

2.6.1. Behavioral analyses

Audio recordings acquired during fMRI scanning were processed to extract the reaction time (RT) for each word. Reaction times, defined as the time between the onset of the pictures and the beginning of speech, were determined with check-vocal (Protopapas, 2007) and by visual inspection of the recordings. Possible differences in the average RTs across conditions were evaluated by using a 2 (active language: German vs. English) $\times$ 2 (trial type: switch vs. stay) repeated-measures ANOVA. All behavioral analyses were performed by using SPSS version 21 for OSX.

2.6.2. Image preprocessing

Preprocessing, parameter estimation, and group statistics of the functional data were performed using SPM12 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) and MATLAB 2013a. Images were realigned and unwarped by using a field map. Estimated movement parameters were checked in order to identify possible excessive head movements. All subjects included in the analyses showed head movements within acceptable limits. Unwarped images were slice-time corrected. The T1 image was coregistered to the mean unwarped EPI image from the same subject. We then spatially normalized the EPI images to the MNI space by using parameter estimates computed on the T1 image. Finally all functional images were smoothed with a Gaussian kernel of 6 mm FWHM.

2.6.3. Neuroimaging analyses

Neuroimaging analyses were performed only on experimental trials, catch trials were not considered. A finite impulse response (FIR) model was applied to the preprocessed images (Henson, 2003). The FIR model was preferred over the standard HRF model in order to better tease apart the bold signal related to the two main task phases: the language intention phase and the language execution phase (see e.g., Reverberi et al., 2012). Each condition was modeled using 16 time bins of 2 s each. We considered two conditions corresponding to trials in which the active language was either German or English. All trials were used for the estimation of the parameters. The onset time for both conditions was the cue appearance. Beside these conditions of interest we also included the movement parameters estimated during realignment as six covariates of no interest in order to remove possible artifacts due to head movement. Linear compounds (contrast) were used to determine the voxelwise time-resolved difference in hemodynamic response between trials cued for English and trials cued for German. In this way 16 contrast images (one per each time bin) were available for each subject. The first ten time bins from each subject were then submitted to a second-level analysis in SPM12 in order to test for the presence of values different from zero. Two main time windows were considered in the analysis. The first was aimed at exploring the activation associated to the representation of the active language and the preparation for its use in the *language intention phase*. The second was aimed at exploring the use of the active language during the *language execution phase*. We assumed an hemodynamic delay of two time-bins from the onset of the cognitive event targeted. Given that, the first two time bins would estimate the bold signal roughly related to the 4 s before cue presentation. Thus, for the intention phase the relevant time bins would be from 3 to 6 (roughly covering 8 s starting from cue presentation), while for the execution phase from time-bin 7 to 8 (roughly covering 4 s during the target phase). The time-bins belonging to each of the two time windows were linearly combined at the second level analysis by means of an *F*-contrast (two-tailed test). In this way we can ensure that no hemodynamic activity belonging to the execution phase influenced the tests for the intention phase. 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 (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.

The univariate analysis on trial sequence was highly similar to that on the active language. The same approach was used both for first and second level analyses. The only difference is that the active language regressor has been substituted by the trial type regressor. Trial type regressor had two level: switch trial vs. stay trial. We considered as switch trials all trials in which the active language has changed from the preceding trials, irrespectively of the specific languages involved.

Finally we also checked for presence of language $\times$ switching interactions. For this analysis a 2 (active language) $\times$ 2 (switch vs. stay trial) first-level model was implemented. For evaluating the presence of an interaction we computed for each subject the contrast (switch–stay)$_{English}$–(switch–stay)$_{German}$ time bin by time bin. In this way we obtained 16 images for each subject. As for the other analyses, the first 10 images were used for the second-level between subject analysis. The presence of an interaction was evaluated in both language intention and execution phases, by means of a two tailed *F* test.

2.6.4. Region of Interest (ROI) analyses

We ran a Region of Interest (ROI) analysis to further assess the involvement of those brain regions found in the whole brain analyses. Specifically, the ROI analysis had two main goals: the first was to assess whether brain regions that showed an effect in one task

phase (e.g., naming an object in a particular language) would also show an effect in the other (e.g., maintaining the intention to use a particular language). The second goal was to determine whether brain regions that showed an effect in the whole-brain language analysis would also show a switching effect, and vice versa whether brain regions showing a switching effect in the whole-brain analysis would also show a language effect. The ROIs were built using Marsbar toolbox (Brett, Anton, Valabregue, & Poline, 2002), based on the results of the whole-brain analyses. All voxels belonging to the same cluster as identified by the automatic procedure in SPM12 were assigned to the same ROI. The activation in the target contrast of all voxels belonging to the same ROI was averaged. Bonferroni correction was used to correct for multiple comparison when multiple ROIs were identified. All ROI analyses were devised so that the definition of the ROI was always independent from the specific comparison performed.

3. Results

3.1. Behavioral results

All subjects selected for fMRI scanning spoke German as first and English as their second language. The average reported proficiency in LEAP-Q for reading, listening and speaking in English was 8.5 (SD = .8), while for German it was 9.8 (SD = .5) on a scale from 1 to 10. The average age of first exposure to English was 9.1 years (SD = 2.3) and .5 years (SD = .8) for German. The average age of proficiently speaking English was 14.7 (SD = 2.7), for German 3.2 (SD = 1). Subjects showed high proficiency in the naming task administered outside the scanner. The overall accuracy was on average 97% (SD = 4%, range: 0–11%), and it was high for both German (M = 98%, SD = 4%, range: 0–11%) and English (M = 96%, SD = 5%, range: 0–13%). Because of a technical problem affecting audio recording during fMRI data acquisition, accuracy data of the naming task performed in the scanner were not fully available and thus were not analyzed.

The average RT across subjects was 1446 ms (SD = 258 ms) for trials in which German was cued, and 1427 ms (SD = 251 ms) for English. Furthermore we also evaluated RTs for trials in which a change of the cued language occurred (switch trials German to English: 1420 ms, SD = 263 ms; switch trials English to German: 1467 ms, SD = 259) and those in which it did not (stay trials English: 1440 ms, SD = 248 ms; stay trials German 1433 ms, SD = 262). In order to check the presence of any reliable difference in reaction times across conditions we performed a 2 (language) × 2 (switch vs. stay) ANOVA. Neither the main effect of language ($F(1,20)$ = 1.378, p = .254) or switch ($F(1,20)$ = .339, p = .567) was significant. The interaction language × switch was not significant ($F(1,20)$ = 3.594, p = .073). Note that the frequently reported effect of switching on reaction times is less likely to show up in our experimental design since the subjects had a delay period of 8 s to prepare prior to speaking.

3.2. Neuroimaging analyses on active language

In this set of analyses we explored whether any brain region showed a different activation for German vs. English. We considered two task phases: in the first phase (language intention phase, see Fig. 1) subjects are instructed by a cue to represent and maintain the intention to use a particular language over a delay period. In the second phase (language execution phase), subjects name the presented object either in German (L1) or English (L2). We ran a time-resolved FIR analysis comparing the relative activation between English and German starting from language cue presentation to language execution.

During the language intention phase we did not find any activation difference between German and English. No difference was found even when the same analysis was performed only for switch trials, or only for stay trials.

A different picture emerged when we considered the execution phase (that is, when the actual target picture appeared). During the execution phase, several brain regions were differentially engaged when subjects were required to name an object in German as compared to English (Fig. 2 and Table 1). More in detail, we found that the brain areas more active when using German (L1) as compared to English (L2) were bilaterally the angular gyrus (Brodmann Area 39), the precuneus, the posterior cingulate cortex, and the right lateral prefrontal cortex (BA 8). By contrast, the brain regions more active when naming in English as compared to German were the inferior frontal gyrus (BA 47) and the insula bilaterally, the anterior cingulate cortex bilaterally (BA 24 and 32), the thalamus and the caudate nucleus bilaterally.

3.3. Neuroimaging analyses on trial sequence

In this set of analyses we compared the trials in which the cued language changed from the preceding trial (switch trials) to the trials in which the cued language remained the same (stay trials). This comparison should detect all brain structures involved in updating the active language. We ran a time-resolved FIR analysis comparing the relative activation between switch and stay trials in the time period from cue presentation to target execution.

During the language intention phase we found a set of posterior brain regions more active in switch compared to stay trials, irrespectively of language (Fig. 3 and Table 2). The active regions were precuneus bilaterally (BA 5 and 7), right superior lateral parietal lobe (BA 7) and the left middle temporal gyrus (BA 39). During the language execution phase we found a higher activation in switch trials only in a small cluster in the prefrontal cortex bilaterally (BA 10).

We finally checked whether the observed activations in switch trials were different across languages. We did not find any brain region showing a significant interaction between switch and language, meaning that the effect of switching is similar across the two languages.

3.4. ROI analyses

We ran Region of Interest (ROI) analyses to further assess whether (i) brain regions showing an effect in one task phase (e.g., language execution) would also show an effect during the other task phase (e.g., language intention); (ii) any brain region showing an effect in whole-brain language analysis would show a switching effect, and vice versa any brain region showing a switching effect would show a different activation depending on the active language. The ROIs were built based on the whole-brain results of the previous analyses.

We found that none of the brain regions (Table 1) that showed a different activation between English and German during language execution showed an effect during language intention. This further confirms the negative finding in the whole-brain analysis on the language effect during intention phase.

The overall picture changed when the switching effect was considered. The fronto-polar ROI that showed a switching effect during language execution (Table 3) also showed the same effect during language intention (p < .05). Furthermore, out of the three ROIs that showed a switching effect during language intention (Table 2), two showed a switching effect also during execution, the superior parietal lobe and the middle temporal gyrus (all p < .01, corrected). Note, though, that the latter finding may just be the tail of the hemodynamic response that occurred in these
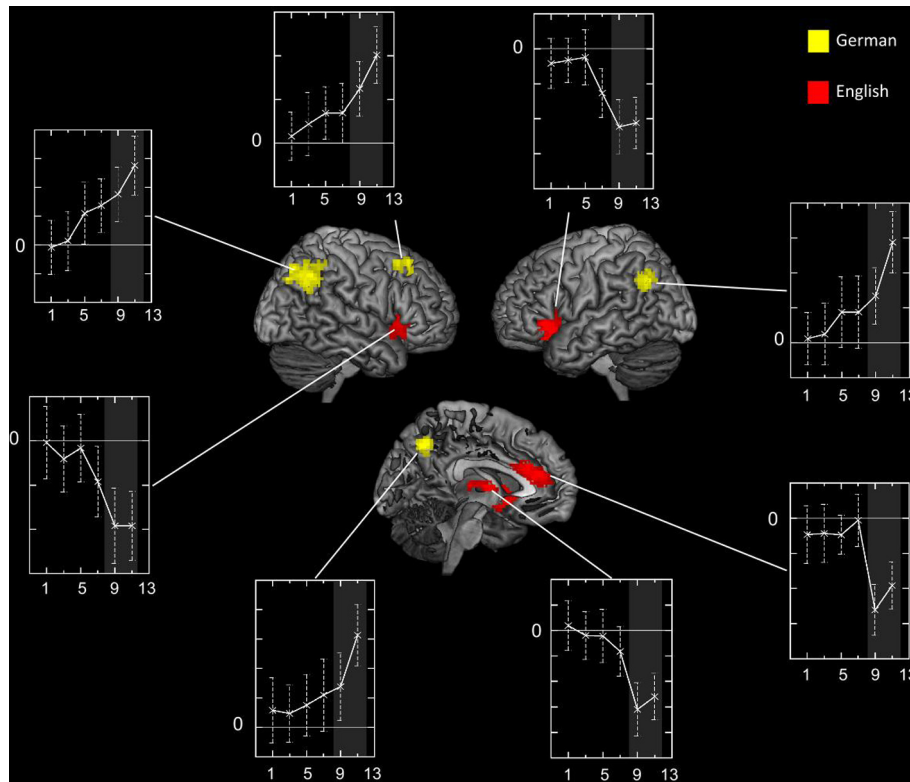
Fig. 2. Activation differences between languages during language execution. The regions that are more activated for English during execution are shown in red and those that are more activated for German are shown in yellow. The corresponding FIR plots of each region are shown next to the regions. The x-axis reports the timing (s) of the experimental trial (cue onset = 0) roughly corresponding to the plotted FIR regressors. The plots cover the two main trial phases considered: language maintenance phase and the language execution phase. The y-axis is the contrast value of German minus English. The time bins belonging to the execution phase are highlighted.

Table 1
Brain activation related to language use during language execution phase. Comparisons of brain activation for naming in German vs. naming in English.

	Brodmann Area	x	y	z	More active language	Z score	Cluster size (k)
Insula		42	20	−1	Eng	4.17	118
Inferior Frontal Gyrus	47	48	23	−10	Eng	3.99	
Insula		33	26	−4	Eng	3.8	
Insula		−33	20	−7	Eng	5.32	190
Inferior Frontal Gyrus	47	−36	29	−4	Eng	5.16	
Insula		−30	14	−13	Eng	4.07	
Anterior Cingulate	32	−3	29	23	Eng	5.1	229
Anterior Cingulate	24	3	32	14	Eng	4.99	
Anterior Cingulate	32	−3	38	20	Eng	4.93	
Thalamus		0	−4	8	Eng	5.6	245
Thalamus		0	−19	8	Eng	4.46	
Caudate		12	17	−4	Eng	4.34	
Superior Frontal Gyrus	8	24	29	53	Ger	4.42	67
Middle Frontal Gyrus	8	30	20	53	Ger	4.21	
Inferior Parietal Lobule	40	57	−55	41	Ger	4.77	195
Angular Gyrus	39	39	−70	47	Ger	4.6	
Angular Gyrus	39	42	−76	38	Ger	4.49	
Angular Gyrus	39	−57	−61	32	Ger	4.5	68
Angular Gyrus	39	−51	−55	35	Ger	3.97	
Angular Gyrus	39	−51	−61	44	Ger	3.81	
Precuneus	7	−6	−61	44	Ger	5.34	92
Precuneus	7	6	−58	44	Ger	3.67	
Posterior Cingulate	31	9	−49	32	Ger	3.29	

brain regions during the preceding maintenance phase. To check whether this was indeed the case we run a new control FIR analysis that focused only on the time-bins of the target phase. Critically in this analysis we covaried out signal likely due to the preceding maintenance phase by modeling it with a canonical HRF time

locked to cue presentation and with duration equal to the maintenance phase. We then re-ran the same ROI analysis reported above on the superior parietal lobe and the middle temporal gyrus. This time none of the two ROIs showed a significantly higher effect in switch compared to stay trials ($p > .7$).
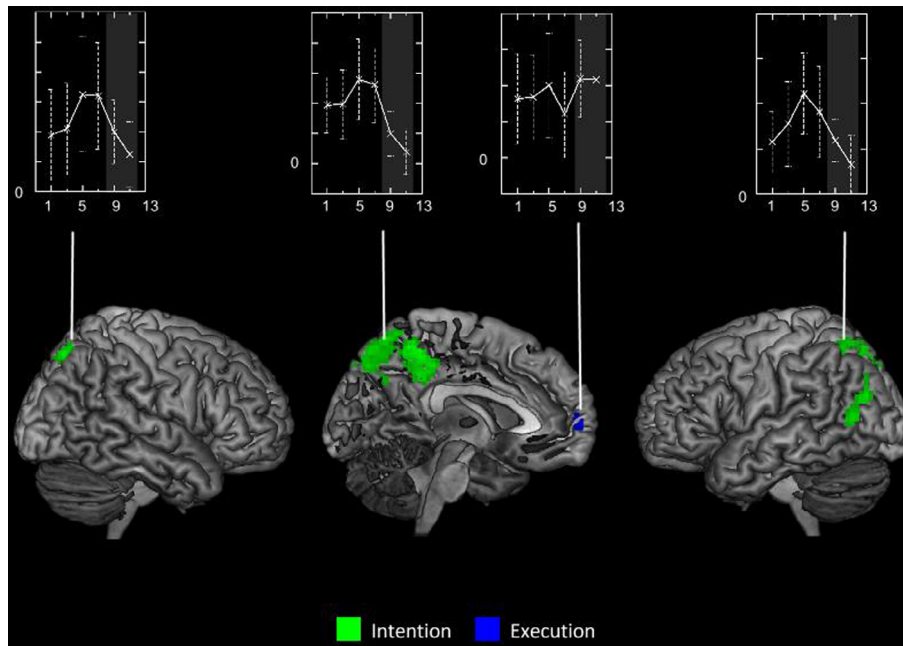
Fig. 3. Activation differences between switch and stay trials during intention and execution phases. The regions shown in green are more activated for switch vs. stay trials during intention phase. The regions shown in blue are more activated for switch trials during execution phase. The corresponding FIR plots of each region are shown next to the regions. The x-axis reports the timing (s) of the experimental trial (cue onset = 0) roughly corresponding to the plotted FIR regressors. The plots cover the two main trial phases considered: language maintenance phase and the language execution phase. The y axis is the contrast value of switch minus stay. The time bins belonging to the execution phase are highlighted.

Table 2
Switch related activation during language intention phase. Brain areas are reported that activated more for switch trials compared to stay trials.

	Brodmann Area	x	y	z	Z score	Cluster size (k)
Precuneus	5	0	−46	56	4.66	447
Cingulate	23	0	−40	41	4.35	
Precuneus	7	−9	−64	62	4.23	
Superior Parietal Lobe	7	21	−73	53	4.38	41
Middle Temporal Lobe	39	−42	−76	23	4.14	74
Middle Temporal Lobe	39	−48	−67	8	4.06	

Table 3
Switch related activation during language execution phase. Brain areas are reported that activated more for switch trials compared to stay trials.

	Brodmann Area	x	y	z	Z score	Cluster size (k)
Medial Frontal Cortex	10	0	62	5	4.19	55
Medial Frontal Cortex	10	9	59	5	4.05	

We then considered whether brain areas that showed a switching effect would also show differences between German and English, either during the intention phase or execution phase. No differences were found (all $p > .5$). Finally, we assessed the presence of switching effects in language ROIs. Only one ROI out of eight showed a switching effect during the intention phase, the one in the Precuneus ($p < .05$, corrected). None of the ROIs showed an effect during execution phase.

4. Discussion

This study aimed at advancing our knowledge of how language control proceeds in bilingual speakers. In particular, we aimed at disentangling the processes involved in language selection (i.e., the "abstract" intention to speak in a given language) from those involved in lexical access (i.e., the actual retrieval of words from a given language). Our study was therefore designed to explicitly tease apart the language intention phase (formulating pre-verbally the intention to speak a particular language while not yet knowing *what* to say) from the language execution phase (naming an object in the intended language). Our results advocate two different systems enabling language selection during bilingual language production. First, a system that it is active during language intention, and that assists in setting up and updating the preparation for speaking a particular language. This system does not appear to rely upon neural areas typically involved in language control (see Abutalebi & Green, 2007). Second, a system active during language execution, which is likely to be in charge of resolving competition between languages when language-specific lexical items need to be selected. This system relies upon the brain network generally reported to be involved in language control in bilinguals. We discuss our findings in detail in the following.

During the language intention phase distinct neural activity is associated with trial sequences in which speakers switched between two languages compared to trial sequences in which speakers stayed with the same language. Switch trials show greater activation in the precuneus bilaterally (BA 5 and 7), in the right superior lateral parietal lobule (BA 7), and in the left middle temporal gyrus (BA 39) when formulating the intention to use a particular language compared to stay trials. Switch trials demand the speaker to re-select the language that will be produced. These trials are therefore assumed to require more cognitive control compared to those trials in which speakers use the same language consecutively, supposedly to override the language previously used (Abutalebi & Green, 2008; Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Gollan & Ferreira, 2009; Meuter & Allport, 1999). In the context of our study, the observed activation differences indicate that, independent of which language will be spoken, the system needs to set-up for

speaking a language different from the one used previously. This interpretation is consistent with other studies that have linked the precuneus to a fronto-parietal control network responsible for directing selective attention (Dosenbach et al., 2007; Utevsky, Smith, & Huettel, 2014), and to studies showing the involvement of the precuneus in selecting and maintaining the intention of performing specific future tasks (prospective intentions, Burgess, Gonen-Yaacovi, & Volle, 2011; Momennejad & Haynes, 2013; Schacter, Addis, & Buckner, 2007; Soon, He, Bode, & Haynes, 2013).

During the language execution phase we find distinct neural activity depending on which language is spoken. Specifically, we observe activation in brain areas that have been associated with cognitive control exerted during bilingual language production, most notably the anterior cingulate and the caudate (Abutalebi & Green, 2007; Abutalebi et al., 2013). These areas were more active when speaking English (L2) compared to German (L1), suggesting that speaking in a second language may require greater cognitive control than speaking in a first language (Garbin et al., 2011). In detail, Abutalebi et al. (2013) report that both the left caudate and the ACC respond differently to different languages in multilinguals. According to this study, the activation in these two areas is higher for the language with less proficiency, typically the L2 or L3. Our findings are therefore in line with the idea that control processes regulate the use of one language over the other in bilingual speech production.

This conclusion is further supported by our analysis comparing switch trials with stay trials. While differences between switch and stay trials are likely to be less pronounced during the language execution phase (compared to the language intention phase discussed above) given the prior long intention phase, nevertheless, we find that switch trials also show a slightly different pattern of activation than stay trials during the language execution phase. Crucially, the observed difference is present in the medial frontal cortex bilaterally. This area is tightly related to cognitive control in bilingual language production. This finding thus supports previous research suggesting that control processes are active in late stages of language production, possibly as late as articulation (e.g., Kroll et al., 2006).

Clearly, not all activation differences we find must be related to cognitive control. In fact, next to lexical access, naming an object is likely to involve many different processes, for example articulatory planning and motor execution. When comparing activation patterns between German and English, such differences could, in principle, be linked to specific features of the two languages involved. However, to date research on the neural basis of bilingualism has not been able to confirm the existence of language-specific differences (for review see Abutalebi & Green, 2007; Abutalebi, 2008), with the possible exceptions of languages such as Chinese as compared to European languages (more right hemispheric activity for the former, see Perfetti et al., 2007). We are therefore confident that our results are not influenced by any putative language-specific differences.

When comparing English (L2) to German (L1) during the language execution phase we find a pattern of brain activity that corresponds to the engagement of the language control network described by Abutalebi and Green (2007; see also Green & Abutalebi, 2013). In other words, during the execution phase, the engagement of this network emphasizes that L2 necessitates more control processes compared to L1. On the other hand, for the reverse comparison, German vs. English, we find that German (L1) activates more bilaterally the inferior parietal lobules, the precuneus, the posterior cingulate cortex, and the right lateral prefrontal cortex. This network is not typically linked to language control but rather to the so-called default mode network (Zhang & Li, 2012). The exact functions of the default mode network remains still to be fully understood; however, it has been recently reported that activity in the default mode network negatively correlates with activity in regions involved in attention and executive function (Fornito, Harrison, Zalesky, & Simons, 2012).

Interestingly, we do not find differences in neural activation during the language intention phase when comparing the two languages. Note that at this point speakers know *which* language they will speak, but they do not yet know *what* they will be uttering. Thus, while speakers have formed the pragmatic intention to use a particular language, they have not yet attempted to access any lexical representation. Even when analyzing only trial sequences in which speakers switched from using one language to another (these trials presumably present a stronger manipulation as they require a reconfiguration of task sets) we do not find significant differences between the two languages. The lack of detectable differences in neural activation associated with the intention to use one or the other language suggests that there is no or very little language competition at the preverbal stage. This finding is in line with psycholinguistic accounts that propose that both languages remain active even when the speaker has formed the intention to use one and not the other language (e.g., Green, 1986, 1998; Kroll et al., 2006; Poulisse & Bongaerts, 1994).

Consistent with proposals by other researchers who have argued for a flexible architecture supporting language selection (Abutalebi & Green, 2007; Kroll et al., 2006), our findings suggest that there may not be a single mechanism or brain structure responsible for selecting one language over another. Instead, different brain regions may be responsible for forming (and maintaining) the intention to speak a particular language, and for executing this intention. During language intention the demands on cognitive control are of a more general nature and do not appear to be specific to the selected language. Hence, we find no activation differences when comparing the two languages directly. Most importantly, during the intention phase we do not find activation in areas related to cognitive control during language execution ("language control") such as the ACC, left caudate, prefrontal cortex and inferior parietal lobule (Green & Abutalebi, 2013). This negative finding makes sense when considering the functional role attributed to these regions, namely the monitoring of the eventual choice of a specific language, signaling errors, and inhibiting the unintended language. These processes only become relevant at the execution phase. In conclusion, our negative finding is not compatible with theories proposing that the unintended language is suppressed already at a preverbal stage, when the intention to speak a particular language is made (La Heij, 2005). On the contrary, our findings suggest that during the intention stage the cognitive system only maintains the abstract intention to use one specific language without pre-activating those regions involved in executing this intention. During language execution, on the other hand, the demands on language control are specific to the language spoken and, consequently, such demands can be influenced by language proficiency. This finding is in line with literature that suggests the cognitive control necessary to resolve competition between languages is only present at the lexical level, i.e., when a specific lexical representation needs to be retrieved in the face of competing alternatives such as translation equivalents (e.g., Abutalebi & Green, 2007).

In summary, implementing an experimental design that explicitly teases apart language selection from lexical access allowed us to disentangle two different networks responsible for language selection that are based on largely different brain structure and that act upon different representations.

Acknowledgments

References

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex, 49*(3), 905–911.

Abutalebi, J., Annoni, J. M., Seghier, M., Zimine, I., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., & Khateb, A. (2008). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cerebral Cortex, 18*, 1496–1505.

Abutalebi, J. (2008). Neural processing of second language representation and control. *Acta Psychologica, 128*, 466–478.

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

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes, 23*(4), 557–582.

Brass, M., & von Cramon, D. Y. (2002). The role of the frontal cortex in task preparation. *Cerebral Cortex, 12*(9), 908–914.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox. *NeuroImage, 16*.

Burgess, P. W., Gonen-Yaacovi, G., & Volle, E. (2011). Functional neuroimaging studies of prospective memory: What have we learnt so far? *Neuropsychologia, 49*(8), 2246–2257.

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish bilinguals. *Bilingualism: Language and Cognition, 2*(03), 231–244.

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.

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.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(5), 1057–1074.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., et al. (2006). Language control in the bilingual brain. *Science, 312*(5779), 1537–1540.

De Groot, A. M. B. (1993). Word-type effects in bilingual processing tasks: Support for a mixed-representational system. In R. Schreuder & B. Weltens (Eds.), *Studies in bilingualism* (Vol. 6, pp. 27). Amsterdam: John Benjamins Publishing Company.

Dosenbach, N. U. F., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A. T., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences, 104*(26), 11073–11078.

Dosenbach, N. U. F., Visscher, K. M., Palmer, E. D., Miezin, F. M., Wenger, K. K., Kang, H. C., et al. (2006). A core system for the implementation of task sets. *Neuron, 50*(5), 799–812.

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition, 9*(02), 153–166.

Fornito, A., Harrison, B. J., Zalesky, A., & Simons, J. S. (2012). Competitive and cooperative dynamics of large-scale brain functional networks supporting recollection. *Proceedings of the National Academy of Sciences, 109*(31), 12788–12793.

Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin, 125*(2), 193–222.

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), 223–235.

Garbin, G., Costa, A., Sanjuan, A., Forn, C., Rodriguez-Pujadas, A., Ventura, N., et al. (2011). Neural bases of language switching in high and early proficient bilinguals. *Brain and Language, 119*(3), 129–135.

Gehrig, J., Wibral, M., Arnold, C., & Kell, C. A. (2012). Setting up the speech production network: How oscillations contribute to lateralized information routing. *Frontiers in Psychology, 3*, 169.

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.

Green, D. W. (1986). Control, activation, and resource. A framework and a model for the control of speech in bilinguals. *Brain and Language, 27*(2), 210–223.

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

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

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism.* Harvard University Press.

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), 676–687.

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, et al. (Eds.), *Human brain function* (2nd ed.) Academic Press.

Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish-English bilinguals: An fMRI study. *NeuroImage, 14*(2), 510–520.

Kell, C. A., Morillon, B., Kouneiher, F., & Giraud, A.-L. (2011). Lateralization of speech production starts in sensory cortices—A possible sensory origin of cerebral left dominance for speech. *Cerebral Cortex, 21*(4), 932–937.

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.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 289–307). New York: Oxford University Press.

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes, 27*(10), 1479–1488.

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.

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.

Momennejad, I., & Haynes, J.-D. (2013). Encoding of prospective tasks in the human prefrontal cortex under varying task loads. *The Journal of Neuroscience, 33*(44), 17342–17349.

Perfetti, C. A., Liu, Y., Fiez, J., Nelson, J., Bolger, D. J., & Tan, L.-H. (2007). Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Bilingualism: Language and Cognition, 10*(02), 131–146.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15*(1), 36–57.

Price, C. J., Green, D. W., & von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain, 122*(12), 2221–2235.

Protopapas, A. (2007). Check vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*(4), 859–862.

Reverberi, C., Görgen, K., & Haynes, J.-D. (2012a). Compositionality of rule representations in human prefrontal cortex. *Cerebral Cortex, 22*(6), 1237–1246 (New York, N.Y.: 1991).

Reverberi, C., Görgen, K., & Haynes, J.-D. (2012b). Distributed representations of rule identity and rule order in human frontal cortex and striatum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32*(48), 17420–17430.

Rodriguez-Fornells, A., van der Lugt, A., Rotte, M., Britti, B., Heinze, H.-J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience, 17*(3), 422–433.

Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb–particle combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(4), 904–921.

Sakai, K., & Passingham, R. E. (2003). Prefrontal interactions reflect future task operations. *Nature Neuroscience, 6*(1), 75–81.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience, 8*(9), 657–661.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*(2), 197–212.

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*.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247–250.

Utevsky, A. V., Smith, D. V., & Huettel, S. A. (2014). Precuneus is a functional core of the default-mode network. *The Journal of Neuroscience, 34*(3), 932–940.

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.

Zhang, S., & Li, C. R. (2012). Functional connectivity mapping of the human precuneus by resting state fMRI. *NeuroImage, 59*(4), 3548–3562.