Behavioral/Cognitive

Bilingual Language Control in Perception versus Action: MEG Reveals Comprehension Control Mechanisms in Anterior Cingulate Cortex and Domain-General Control of Production in Dorsolateral Prefrontal Cortex

Esti Blanco-Elorrieta[1,2] and Liina Pylkkänen[1,2,3]

[1]New York University Abu Dhabi Institute, Abu Dhabi, United Arab Emirates, and [2]Department of Psychology and [3]Department of Linguistics, New York University, New York, New York 10003

For multilingual individuals, adaptive goal-directed behavior as enabled by cognitive control includes the management of two or more languages. This work used magnetoencephalography (MEG) to investigate the degree of neural overlap between language control and domain-general cognitive control both in action and perception. Highly proficient Arabic–English bilingual individuals participated in maximally parallel language-switching tasks in production and comprehension as well as in analogous tasks in which, instead of the used language, the semantic category of the comprehended/produced word changed. Our results indicated a clear dissociation of language control mechanisms in production versus comprehension. Language-switching in production recruited dorsolateral prefrontal regions bilaterally and, importantly, these regions were similarly recruited by category-switching. Conversely, effects of language-switching in comprehension were observed in the anterior cingulate cortex and were not shared by category-switching. These results suggest that bilingual individuals rely on adaptive language control strategies and that the neural involvement during language-switching could be extensively influenced by whether the switch is active (e.g., in production) or passive (e.g., in comprehension). In addition, these results support that humans require high-level cognitive control to switch languages in production, but the comprehension of language switches recruits a distinct neural circuitry. The use of MEG enabled us to obtain the first characterization of the spatiotemporal profile of these effects, establishing that switching processes begin ~400 ms after stimulus presentation.

Key words: adaptive cognitive control; bilingualism; comprehension; language control; magnetoencephalography (MEG); production

Significance Statement

This research addresses the neural mechanisms underlying multilingual individuals' ability to successfully manage two or more languages, critically targeting whether language control is uniform across linguistic domains (production and comprehension) and whether it is a subdomain of general cognitive control. The results showed that language production and comprehension rely on different networks: whereas language control in production recruited domain-general networks, the brain bases of switching during comprehension seemed language specific. Therefore, the crucial assumption of the bilingual advantage hypothesis, that there is a close relationship between language control and general cognitive control, seems to only hold during production.

Introduction

Our ability to organize thoughts and actions compliant with internally defined goals is commonly referred to as cognitive con-

trol (Miller, 2000). For multilingual individuals, comprising approximately half of the world's population, this capacity includes the management of two or more languages, exemplified by these individuals' ability to communicate in one language with no obvious lexical or syntactic interference from the other. Therefore, a critical question for understanding the neural architecture of the bilingual brain is whether bilingual language control is part of general cognitive control (Craik and Bialystok,

2006; Garbin et al., 2010; Abutalebi et al., 2012) or if bilingual individuals develop specialized mechanisms to control language (Abutalebi et al., 2008; Calabria et al., 2011). The controversial bilingual advantage hypothesis (de Bruin et al., 2015) critically assumes a close relationship between language control and general cognitive control, making this question relevant, not only for the neuroscience of bilingualism, but also for educational planning.

Because language-switching occurs in comprehension and production, the question about the domain generality of language control interacts with another unaddressed basic question about the neurobiology of bilingualism: does language-switching in comprehension and production use similar brain mechanisms? The predominant bilingual language comprehension model (van Heuven and Dijkstra, 2010) predicts so, because it proposes that comprehension and production recruit similar language-external inhibitory control mechanisms (see also Abutalebi et al., 2007). However, the disposition to switch clearly varies in these two domains, with production switches being consciously decided upon and comprehension switches inertly perceived, which could, by hypothesis, lead to differences in the recruited networks.

Prior language control research has mostly focused on production (Hernandez et al., 2001; Rodriguez-Fornells et al., 2005; Wang et al., 2007; Guo et al., 2011), leaving the neural underpinnings of language-switching in comprehension rather elusive (but see Abutalebi et al., 2007). In addition, the direct contrasts (i.e., same population tested on similar tasks involving language and general executive control) conducted in production have so far not reached a unanimous verdict regarding the degree of anatomical overlap between high-level cognitive control and bilingual language control: whereas some studies propose a nearly complete overlap between the two (De Baene et al., 2015; Weissberger et al., 2015), others suggest only partial overlap (Branzi et al., 2015) or mostly distinct networks (Abutalebi et al., 2008; Magezi et al., 2012). Here, we targeted both questions within the same design, addressing not only the neurobiological similarity of language-switching and non-language-switching, but also the commonality between language-switching in comprehension and production.

We designed maximally parallel language-switching tasks in production/comprehension and compared the results with analogous tasks in which, instead of the used language, the semantic category of the comprehended/produced word changed. In both production tasks, participants named playing cards, the color of the suit cueing output selection. In language-switching, participants named the numerosity depicted by the card, red suits standing for Arabic and black ones for English. In category-switching, performed in Arabic, red indicated numerosity naming and black suit naming (Fig. 1). In the language-switching version of the comprehension tasks, subjects listened to number words in Arabic/English and indicated whether a subsequent visually presented colored number matched the auditory input. In the category-switching variant, subjects heard number/color words in Arabic and indicated whether the visually presented number matched what they had heard. Therefore, we sought maximal similarity both across modalities (comprehension/production) and domains (language/non-language-switching), making direct comparisons between all conditions maximally informative.

Brain activity was monitored millisecond by millisecond with magnetoencephalography (MEG) and analyzed in ROIs defined based on prior research on general cognitive control (Hikosaka

and Isoda, 2010; Braver et al., 2001), language-switching in production (Hernandez et al., 2000, 2001; Rodriguez-Fornells et al., 2005; Crinion et al., 2006; Abutalebi et al., 2008), and language-switching in comprehension (Abutalebi et al., 2007). Our goal was to first characterize the degree of neural overlap between language-switching in production versus comprehension and then to assess the extent to which such switching effects are domain general.

Materials and Methods

Participants
Nineteen right-handed bilingual individuals participated in the experiment (14 male, 5 female, M = 20.2 years, SD = 2.6). All were native speakers of Arabic with a high knowledge of English (M = 7.75 in a 1–10 scale, SD = 1.17). They all came from Arab families but had always lived in an Arabic–English bilingual environment [M (exposure to Arabic) = 45.42%, M (exposure to English) = 54.58%, SD = 18.85] and were enrolled in an English-speaking university. Mean ages of acquisition were 0.88 (SD = 0.56) for Arabic and 1.2 (SD = 0.56) for English. Information about their language use and proficiency level was gathered with a language background questionnaire (Marian et al., 2007). All subjects were neurologically intact with normal or corrected-to-normal vision and all provided informed written consent.
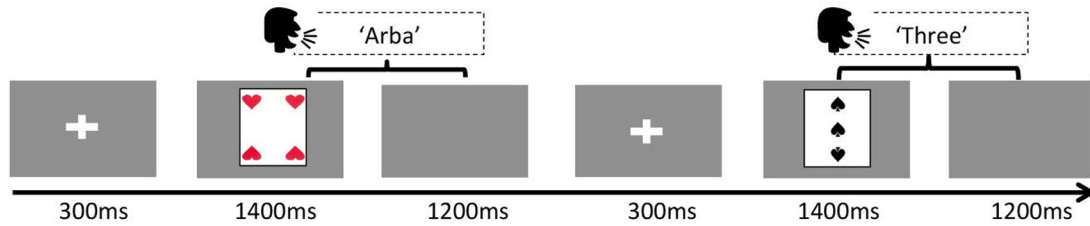
Stimuli and experimental design
The experiment consisted of four tasks: (1) language-switching in production, (2) category-switching in production, (3) language-switching in comprehension, and (4) category-switching in comprehension. In both production tasks, participants named playing cards. Crucially, selecting playing cards as stimuli allowed us to create different naming tasks while keeping the visual stimulus constant. The numerosity displayed on the card only varied from one to four because four has been established as the upper bound for responses to be subitized (Saltzman and Gamer, 1948; Kaufman et al., 1949). In addition, because in category-switching, participants had to name one of four different suits, using only four numbers allowed for the number of stimuli to be kept constant across both production tasks (see Fig. 1 for design and examples).

Further, within language-switching, the target words for production constituted the auditory stimuli in the comprehension task. In the category-switching tasks, number words constituted one of the stimulus categories both in production and comprehension; however, suit names were the other stimulus category in production and color words in comprehension. The reason is that there is significant variation in suit names among Arabic dialects and, although this was not a problem in production (participants were asked to name them in the most natural way for them), it presented a problem in comprehension (participants may not have recognized the auditory stimulus if presented with a suit name from a different dialect). Therefore, category-switching in comprehension encompassed switches between numbers and color words. Specifically, we used red (aHmar احمر), black (aswad اسود), blue (azraq ازرق), and green (axDar اخضر). An Arabic–English bilingual speaker recorded all auditory stimuli in a single session using a Neumeann U87 microphone and Avalon VT-737SP preamplifier. The speaker read each word three times and the second production of the word was always selected to allow for consistent intonation across stimuli.

The experiment consisted of a total of 768 trials. Half belonged to the production and half to the comprehension tasks and were evenly distributed across the language and category-switching tasks. Therefore, each of the four experimental conditions contained 192 items. These items were presented in one of three types of trials: (1) trials in which the language/semantic category of the target item differed from that of the preceding trial (switch trials), (2) trials in which the language/semantic category of the target was identical to that of the preceding trial but immediately followed a switch trial (switch + 1 trials), and (3) trials in which the targets' language/category was identical to that of the preceding trial and did not follow a switch trial (non-switch trials). Some previous evidence has shown that switch costs are not always confined to the first trial after a switch (Salthouse et al., 1998; Meiran et al., 2000); therefore, "switch +
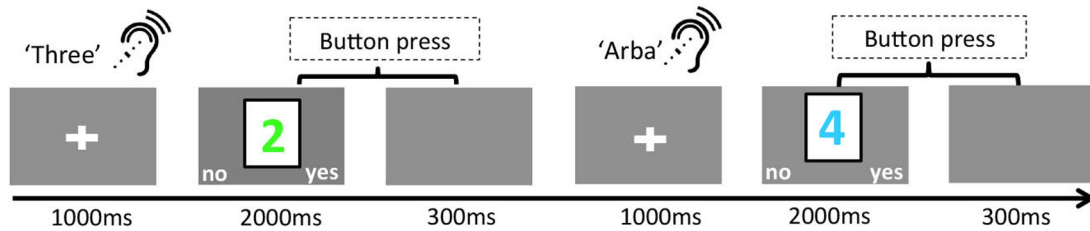
Figure 1. Experimental design. In two of the tasks (**A**, **C**), participants were asked to describe the playing cards presented onscreen by (**A**) naming the numerosity depicted by the card (red suits standing for Arabic and black ones for English) or (**C**) naming the numerosity or the suit depicted by the card, both of them in Arabic. In the other two tasks (**B**, **D**), participants were asked to listen to an auditory stimulus and judge via button press whether a subsequent visually present colored number matched the auditory input. In **B**, participants listened to number words in Arabic or English and, in **D**, participants listened to number and color words in Arabic.
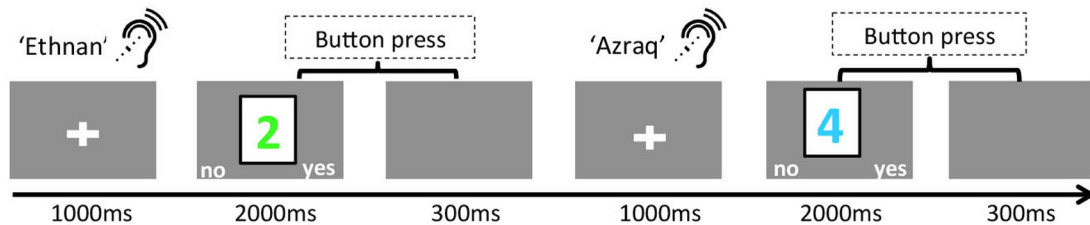
1" and "non-switch" trials were distinguished by hypothesizing that "switch + 1" could carry over some cost of the switch (Allport et al., 1994), so "non-switch" trials may constitute a better baseline with which to measure the switch effect. Importantly, language switches came in an unpredictable fashion.

All trial types contained equal number of repetitions of each experimental item. Experimental conditions were further divided into eight blocks of 24 items to form experimental blocks. Items within blocks and blocks along the experiment were pseudorandomized following two constraints: two identical stimuli were never presented consecutively and two blocks of the same experimental condition never appeared successively. Participants were informed at the beginning of each block whether the upcoming block was going to be a language- or a category-switching block and whether the task was production or comprehension. All pic-

tures were presented foveally using Presentation software (Neurobehavioral System) and subtended in a range from 1.65° height and 2.55° width on a screen ~85 cm from the subject. The size of the picture ensured that only one fixation was required to perceive all of the elements of the stimuli, which was crucial to avoid saccade-related artifacts.

Procedure
Before recording, each subject's head shape was digitized using a Polhemus dual source handheld FastSCAN laser scanner. MEG data were collected in the Neuroscience of Language Laboratory in NYU Abu Dhabi using a whole-head 208 channel axial gradiometer system (Kanazawa Institute of Technology, Kanazawa, Japan) as subjects lay in a dimly lit, magnetically shielded room. Vocal responses were captured with an MEG-compatible microphone (Shure PG 81). In the production

conditions, trials began with a fixation cross (300 ms), followed by the presentation of the picture stimuli. Pictures remained onscreen until speech onset and participants were allowed 1400 ms to respond. Participants were then given 1200 ms to finish speech before the next trial began and they were instructed to respond as quickly and accurately as possible. MEG data were recorded during planning for production, before motion artifacts, and participants were allowed to blink after naming the stimulus.

In the comprehension tasks, participants were presented with a fixation cross while they heard the auditory stimuli. This cross stayed onscreen for ~400 ms after the end of the stimulus (maximum: 420 ms, minimum: 380 ms, due to differences in the length of the auditory stimuli) and this was followed by the presentation of a colored number on the screen. Participants were then given 2000 ms to make a judgment via button press about whether the auditory and the visual stimulus matched. For all participants, the left button indicated mismatch and the right button indicated match. After the button press, a blank screen appeared for 300 ms and then the next trial began (Fig. 1). MEG data were acquired during auditory stimulus presentation to capture activity elicited by perceiving the language switch. The ~400 ms lapse between the end of the auditory stimulus and the presentation of the visual stimulus was included to avoid capturing task-related (i.e., decision making and mismatch detection) activity in our epoch. However, during this lapse, participants likely processed the language switch, so the behavioral reaction times measured from visual stimulus presentation were not informative of online switching. Therefore, only accuracy was considered a reliable measure of participants' switching performance in comprehension tasks. Participants were allowed to blink during visual stimulus presentation, before the button press that initiated the following trial.

Data acquisition and preprocessing

MEG data were recorded at 1000 Hz (200 Hz low-pass filter), noise reduced via the continuously adjusted least-squares method (Adachi et al., 2001) in MEG Laboratory software (Yokogawa Electric and Eagle Technology) and epoched from 200 ms before to 700 ms after critical stimulus onset. Individual epochs were automatically rejected if any sensor value after noise reduction exceeded 2500 fT/cm at any time. Trials containing remaining blinks were identified by individually visualizing raw activity for each epoch. If there was any sudden, stark increase of activity, then the topography for that epoch was plotted. If the magnetic field pattern had the characteristic frontal distribution of a blink, then that trial was rejected. In addition, trials corresponding to behavioral errors or response times within the length of our epochs were also excluded from further analyses. By condition, this resulted in the exclusion of 24.39% of the non-switch, 19.7% of the switch + 1 and 20.92% of the switch trials within the general-domain comprehension task. Within the general-domain production task, 24.47% of the non-switch trials, 23.69% of the switch + 1 trials, and 18.14% of the switch trials were rejected. In the language comprehension task, 19.61% of the non-switch, 19.87% of the switch + 1, and 16.14% of the switch trials were rejected. Finally, within the language production task, 26.9% of the non-switch trials, 27.6% of the switch + 1 trials, and 24.2% of the switch trials were excluded from further analysis. Altogether, this resulted in the exclusion of 25.82% of the trials (SD = 14.46%), leaving 569.63 trials on average per subject (SD = 111.10). To prevent oral artifacts from contaminating our data, we followed the same strict artifact rejection routine that we have followed in our previous production studies (Pylkkänen et al., 2014; Blanco-Elorrieta and Pylkkänen, 2015, 2016): (1) removing all trials that contained naming latencies within our epoch, (2) removing all individual epochs that contained amplitudes >2500 feet/cm for any sensor after noise reduction, (3) visualizing all individual epochs before averaging and rejecting any epoch that contained any sudden increases in the magnitude of the signal caused by artifacts (be it muscular movements or else), and (4) applying a 40 Hz low pass filter that should eliminate any remaining oral movement from our data because the gamma-frequency range (>40 Hz) is reportedly the one affected by muscle artifact contamination such as phasic contractions (Yuval-Greenberg and Deouell, 2009; Gross et al., 2013).

Cortically constrained minimum-norm estimates (Hämäläinen and Ilmoniemi, 1994) were calculated via MNE (MGH/HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging, Charleston, MA). The cortical surfaces were constructed by mapping an average brain from Free-Surfer (CorTech and MGH/HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging) to the head-shape data gathered from the head-scanning process. This generated a source space of 5124 points for each reconstructed surface. Then, the boundary-element model method was used to calculate the forward solution. Epochs were baseline corrected with the pretarget interval [−200 ms, 0 ms] and low-pass filtered at 40 Hz. Using the grand average of all trials for a particular subject, the inverse solution was computed from the forward solution. This determined the most likely distribution of neural activity. The resulting minimum norm estimates of neural activity were transformed into normalized estimates of noise at each spatial location obtaining statistical parametric maps (SPMs), which provide information about the statistical reliability of the estimated signal at each location in the map with millisecond accuracy. Then, those SPMs were converted to dynamic maps (dSPMs). To quantify the spatial resolution of these maps, the point spread function for different locations on the cortical surface was computed, which reflects the spatial blurring of the true activity patterns in the spatiotemporal maps, thus obtaining estimates of brain electrical activity with the best possible spatial and temporal accuracy (Dale et al., 2000).

Data analysis

Behavioral data. In the comprehension tasks, incorrect button presses were coded as errors for accuracy measures. Reaction times were not analyzed because they were elicited after a substantial delay from the stimulus onset and were consequently not informative of online switching (see "Procedure" section). In the production tasks, participants' vocal responses were evaluated for each trial and reaction times corresponding to erroneous responses [incorrect naming, verbal disfluencies (i.e., utterance repairs, stuttering) and nonresponses] were excluded from further analysis. In addition, trials following participants' errors were also excluded when such errors altered the type of subsequent trials (e.g., in a "switch, non-switch" sequence where item 1 was labeled "switch to Arabic" and item 2 "non-switch in Arabic", item 2 was excluded if the participant erred in item 1 by naming it in English, given that this would imply that even if item 2 was correctly named in Arabic, it would not be "non-switch" anymore). Naming latencies below or above 2.5 SD from the mean were also discarded. Reaction times for production and accuracy rates for production and comprehension were averaged over trials per condition and per participant and subjected to 2 × 3 repeated-measures ANOVA [main factors domain (language or category-switching) and switch (switch, switch + 1, non-switch)]. Planned contrasts were also examined with paired *t* tests (two-tailed), applying Bonferroni correction for multiple comparisons. Behavioral switch costs were determined by the differences in naming latencies between switch and non-switch trials within each of the tasks (RT switch$_{language\ production}$ − RT non-switch$_{language\ production}$).

ROI main analyses. Following previous studies that have found language-switching to recruit dorsolateral areas of the prefrontal cortex (dlPFC) in both hemispheres (Hernandez et al., 2000, 2001; Rodriguez-Fornells et al., 2005; Wang et al., 2007), we analyzed Brodmann areas (BAs) 9, 10, and 46 bilaterally. Furthermore, because language-switching has been reported to involve the left ACC (Wang et al., 2007; Abutalebi et al., 2008, Crinion et al., 2006; van Heuven et al., 2008), we also analyzed this area, including BA24, BA32, and BA33 in the analyses (MacDonald et al., 2000). In addition, given that, in the left hemisphere, the prefrontal ROIs neighbored Broca's area, we also ran an analysis to assess whether linguistic switch effects would extend to the left inferior frontal gyrus (LIFG), including in this analysis BA44, BA45, and BA47 (Thompson-Schill et al., 1998; Poldrack et al., 1999; Wagner et al., 2000, 2001; Blanco-Elorrieta and Pylkkänen, 2015). To assign each source to a BA, the Talairach Atlas was used (http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach). Following this atlas, each point in the reconstructed cortical surface (see "Data acquisition and preprocessing" section) was au-

tomatically assigned to the nearest labeled BA and each point in the cerebellum to the nearest gyrus. This assignment of sources to BAs was then compared with an annotation file (lh.aparc.a2009s.annot) from Fs average data of FreeSurfer (CorTech and MGH/HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging). If there was a discrepancy between the two, the sources were manually reassigned following the latter model to a different BA. For each of the BAs and gyri found, an MNE label was created.

The noise-normalized dSPMs (Dale et al., 2000) resulting from the preprocessing of our data were submitted to a nonparametric permutation test to identify temporal clusters with significantly different activation between conditions (Maris and Oostenveld, 2007). This experiment aimed to reveal the loci for language-switching in production and comprehension and the extent to which such regions are also involved in category-switching (but not the opposite). Therefore, in addition to switch effects common to both tasks, only interactions in which there was a difference between language-switch and non-switch conditions, but not between category-switching conditions, were relevant for our purposes. For this reason, following the discussion by Maris and Oostenveld (2007), a test statistic that incorporates this premise and is capable of identifying these particular types of interactions was constructed (see also Bemis and Pylkkänen, 2011).

In the statistical analysis, data for each time point were submitted to 2×2 repeated-measures cluster permutation ANOVAs with domain (language-switching or category-switching) and switch (switch or non-switch) as factors. Clusters were extracted by identifying intervals of at least 10 adjacent time points for which, at each time point, the resulting p-value was less than a set threshold. Following previous MEG language studies, we adopted $p < 0.3$ (uncorrected) threshold for cluster selection, enabling us to capture potential long-lasting effects in the data (Bemis and Pylkkänen 2011, 2012; Del Prato and Pylkkänen 2014; Pylkkänen et al., 2014; Westerlund and Pylkkänen 2014; Blanco-Elorrieta and Pylkkänen 2015, 2016). To calculate the cluster-level statistics of the main effects, F values from the repeated-measures ANOVA were calculated for each time point within the analyzed time windows and used as the test statistic for the permutation tests. To calculate the cluster-level statistics of the interaction, a paired t test was performed at each time point of the analyzed window between trial types in each task. This produced two t values for every time point in the cluster: one corresponding to the difference in activity between the two language conditions (switch and non-switch) and one corresponding to the difference in activity between the two category-switching conditions (switch and non-switch). The absolute value of the second t-value was then subtracted from the first, yielding a component test statistic that was then used in the permutation tests (Bemis and Pylkkänen, 2011). Therefore, the test statistic for interaction clusters was calculated by summing all language-switching t values and subtracting the magnitude of all category-switching t values. Data for both main effects and the interaction were then subjected to 10,000 random permutations and the final corrected p value (α: $p < 0.05$) of the observed data was calculated as the ratio of permutations yielding a test statistic greater than the actual observed test statistic. Clusters with a final corrected p-value <0.1 were considered marginally reliable. Cluster p-values were corrected for multiple comparisons across all BAs entered into a single analysis using the false discovery rate (FDR; $q = 0.05$) (Benjamini and Yekutieli, 2001; Genovese et al., 2002). The permutation tests were conducted in a midlatency (300–500 ms) and a late (500 - 700 ms) window in all analyzed areas. Because permutation tests only report the most significant cluster within an analyzed time window, dividing the analyzed interval was crucial for being able to identify multiple distinctly timed effects. This was particularly important given that previous investigations have not been able to characterize the timing of switching effects due to the lack of time resolution in fMRI.

To unpack the results of the ANOVAs, follow-up two-tailed cluster permutation t tests using parameters identical to the ANOVAs were performed in the same time windows. As an additional confirmatory analysis, we visualized all pairwise effects source by source within our ROIs by plotting each effect that held at least for 15 ms in at least 15 adjacent cortical sources at $p < 0.05$. This allowed for an informal, uncorrected visualization of the spatial extent of our effects within the ROIs.

Additional ROI analyses. To investigate whether the switch effect disappeared gradually after a switch item or if the effect was confined to the first element after a switch, a 1×3 ANOVA was run within each task (main factor switch: switch, switch + 1, Non-switch). Nonparametric permutation ANOVAs were used in these analyses, mirroring the main ROI analysis. However, because we did not hold any particular hypotheses regarding the direction of possible interactions, F values from the repeated-measures ANOVA were calculated for each time point within the analyzed time windows and used in the permutation tests to calculate the cluster-level statistic of the main effects and the interactions.

Sensor data analyses. Finally, we conducted sensor data analyses to test whether any observed effects were accompanied by parallel effects in the MEG sensor data free of source modeling assumptions. The sensor array was divided in quadrants and root mean square calculations of the signal at each sensor were averaged within the partitions and submitted to 2×2 cluster-based permutation ANOVAs identical in all parameters to the ones described above for the ROI analysis. Results in the four partitions were also FDR corrected (Genovese et al., 2002) with a criterion value of 0.05. The aim of our sensor analysis was to provide a simple and easily replicable assessment of our hypotheses, as well as to reduce the gap between our results and the ERP literature. However, the roughness of this analysis, that is, the use of the same sensors across all participants despite different head locations in the helmet, obviously lessened our chances of closely mirroring the source analysis.

Results
Behavioral results
To investigate the nature of behavioral switch effects and to address whether such effects disappear gradually or immediately after a switch, we ran a 2×3 ANOVA on production reaction time data with the factors domain (language or category-switching) and switch (switch, switch + 1, non-switch). This test revealed a significant main effect of domain ($F_{(1,18)} = 106.2$; $p < 0.0001$) and of switch ($F_{(2,36)} = 106.2$; $p < 0.0001$). The interaction between domain and switch was not significant ($F_{(2,36)} = 2.37$; $p = 1.08$). Planned t tests showed that switch + 1 trials were significantly faster than switch trials both in language-switching ($t_{(37)} = -10.48$, $p < 0.0001$) and category-switching ($t_{(37)} = -9.18$, $p < 0.0001$), but reliably slower than non-switch trials both during language-switching ($t_{(37)} = 4.46$, $p = 0.0001$) and category-switching ($t_{(37)} = 3.10$, $p = 0.003$), thus showing that switch effects disappear gradually after a switch item (Fig. 2, Table 1). Note that t tests were only deemed significant if they survived Bonferroni correction for multiple comparisons in which our α was established at 0.0124 (c = α/m; c = 0.05/4 = 0.0124). The switch cost during language production amounted to 111 ms and during category-switching to 145 ms.

Accuracy measures were also submitted to 2×3 ANOVAs in production and comprehension. In production, the ANOVA revealed a reliable main effect of switch ($F_{(2,36)} = 22.36$; $p < 0.0001$), and a reliable interaction between domain and switch ($F_{(2,36)} = 4.59$; $p = 0.01$). The main effect of domain was not reliable ($F_{(1,18)} = 1.08$; $p = 0.3$). Planned t tests corrected with Bonferroni ($\alpha = 0.0124$) showed that participants were significantly more error prone in switch + 1 trials than in non-switch trials both during language-switching ($t_{(18)} = -5.34$; $p < 0.0001$) and category-switching ($t_{(18)} = 4.33$; $p = 0.0003$). There was no reliable difference between switch + 1 and switch trials in any of the domains [language-switching ($t_{(18)} = -0.48$; $p = 0.63$); category-switching ($t_{(18)} = -0.19$; $p = 0.85$)], but switch trials had lower accuracy rates overall (Fig. 2, Table 1). In comprehension, neither the main effect of domain ($F_{(1,18)} = 1.4$; $p = 0.24$) nor the main effect of switch ($F_{(2,36)} = 0.28$; $p = 0.7$) were significant. The interaction between the two ($F_{(2,36)} = 3.12$; $p = 0.05$) did not survive correction for multiple comparisons.
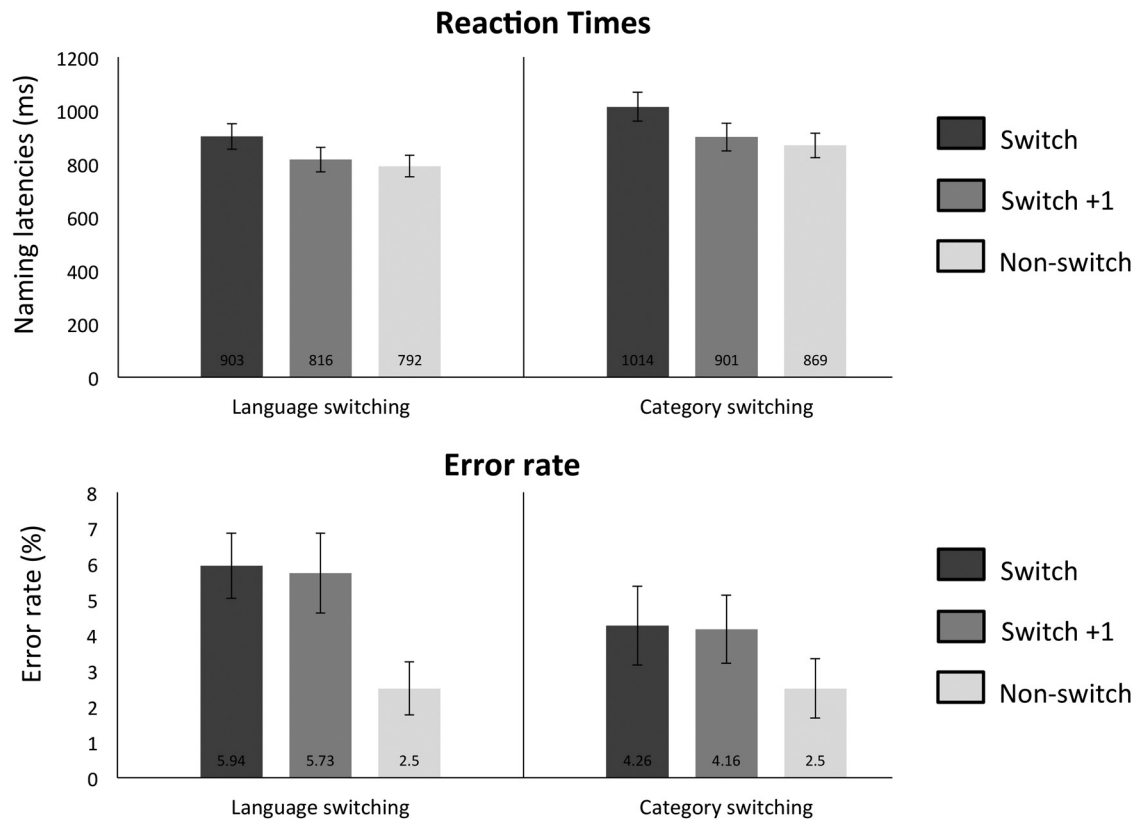
Figure 2. Mean reaction times and error rates as a function of the performed switching condition within production tasks. The number at the bottom of each bar indicates the average value for that condition. Error bars indicate SEM.

Table 1. Mean reaction times and percentage of errors across all conditions

	Switch	Switch + 1	Non-switch
Language-switching			
RT	*903* (48)	*816* (45)	*792* (41)
Error rate	*5.94* (1.1)	*5.73* (1.1)	*2.5* (0.74)
Category-switching			
RT	*1014* (54)	*901* (52)	*869* (46)
Error rate	*4.26* (0.91)	*4.16* (0.95)	*2.5* (0.83)

Means are displayed in italics; SEMs are displayed in parentheses.

ROI results

ROI analyses were performed in the left ACC (BA24, BA32, and BA33), the LIFG (BA44, BA45, and BA47) and in the dlPFC bilaterally (BA9, BA10, and BA46). Non-switch trials were used as the baseline to measure switch effects (see "Stimuli and experimental design"). Therefore, we ran 2 × 2 ANOVAs with domain (category or language-switching) and switch (switch or non-switch) as main factors in a mid (300:500 ms) and late (500:700 ms) time windows in production and comprehension. In addition, a one-way ANOVA was run within each task contrasting the switch, switch + 1, and non-switch conditions to assess whether the predicted higher activity for switch items was constrained to only the first element after a switch or whether it disappeared gradually. To facilitate visualization and interpretation of our analyses, the waveforms for each ANOVA are displayed as two pairwise comparisons in the figures and statistical significance of such pairwise contrasts is shown. In addition, these plots are accompanied by uncorrected visualizations of the spatial extent of our effects within the ROIs.

Switch effects in production

Switch effects for production tasks emerged in the dlPFC. Precisely, the 2 × 2 ANOVA (domain: language-switching/category-switching and switch: switch/non-switch) in the midlatency window revealed a reliable main effect of switch in BA9 (413–500 ms; $p = 0.04$), BA10 (361–500 ms; $p = 0.01$), and BA46 (359–374 ms; $p = 0.02$) in the left hemisphere and in BA9 (346–500 ms; $p = 0.03$), BA10 (350–500 ms; $p = 0.01$), and BA46 (347–500 ms; $p = 0.01$) in the right hemisphere. Planned t tests showed that these main effects of switch were caused by switch trials eliciting increased activity compared with non-switch trials in both tasks (Fig. 3). Precisely, within the language-switching task, switch trials elicited significant or marginally significant increases in activity over non-switch trials in the left BA10 (414–461 ms; $p = 0.09$), right BA9 (601–673 ms; $p = 0.08$), right BA10 (408–500 ms; $p = 0.02$ and 597–678 ms; $p = 0.09$), and right BA46 (412–470 ms; $p = 0.06$ and 500–683 ms; $p = 0.006$). Within category-switching, the clusters for increased activity during switch trials were found in the left BA9 (409–493 ms; $p = 0.02$ and 541–637 ms; $p = 0.07$), left BA10 (391–500 ms; $p = 0.009$), left BA46 (392–500 ms; $p = 0.008$), and right BA9 (350–409 ms; $p = 0.08$). In addition, the ANOVAs conducted on the LIFG and ACC activity did not reveal any reliable cluster either in the earlier (300–500 ms) or in the later (500–700 ms) time windows.

Switch effects in comprehension

In comprehension, the 2 × 2 ANOVA (domain: language-switching/category-switching and switch: switch/non-switch) conducted in the left ACC revealed a reliable interaction between domain and switch in BA24 (575–700 ms; $p = 0.01$) and
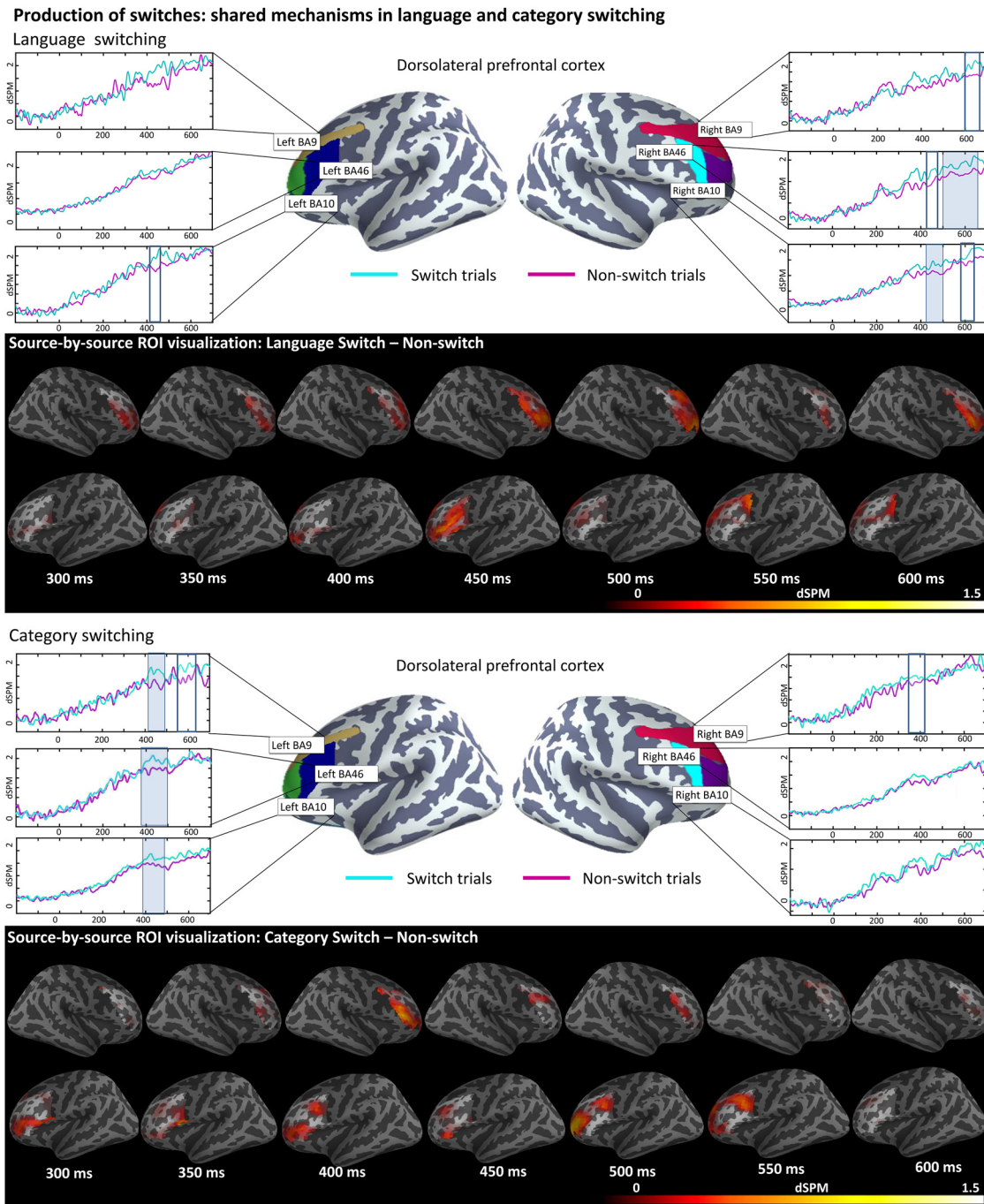
Figure 3. ROI results for pairwise comparisons for production conditions in the dlPFC; activation averaged across subjects. A freeSurfer average brain illustrates the spatial distribution of the BAs included in the analyses color coded by BA. On the waveform plots, the shaded regions indicate that the difference in activity between the two tested conditions was significant at $p = 0.05$ (corrected), whereas the boxed region indicates marginally significant effects at $p < 0.01$ (corrected). Significance was determined using a nonparametric permutation test (Maris and Oostenveld, 2007) performed from 300–500 and 500–700 milliseconds (10,000 permutations). The whole-brain comparisons display all pairwise effects source by source within our ROIs. Only effects in which the first condition (switch condition) elicited greater activity than the second condition (non-switch activity) for at least 15 ms in at least 15 adjacent cortical sources at $p < 0.05$ are displayed, providing a complementary visualization of the spatial extent of our effects within the ROIs. Section A illustrates differences during the language-switching task and section B illustrates comparisons conducted within the category-switching task.

marginally reliable interaction in BA32 (393–424 ms; $p = 0.06$) and BA33 (408–425 ms; $p = 0.09$ and 681–700; $p = 0.07$). This interaction between domain and switch was further unpacked with two-tailed t tests, finding that it was caused by increased activity for switch over non-switch trials within language-switching and no differences between category-switching conditions (Fig. 4). Specifically, contrasting switch trials with non-switch trials within language-switching elicited reliable clusters of activation in BA24 (426–500 ms; $p = 0.05$) and BA32 (398–472 ms; $p = 0.02$) in the earlier time window and in BA24 (551–673 ms; $p = 0.01$) and BA33 (620–700 ms; $p = 0.05$) in the later time window. The t tests within category-switching did not reveal any reliable difference between switch and non-switch trials. The analyses conducted on LIFG and dlPFC activity did not elicit any reliable cluster at $p < 0.05$.
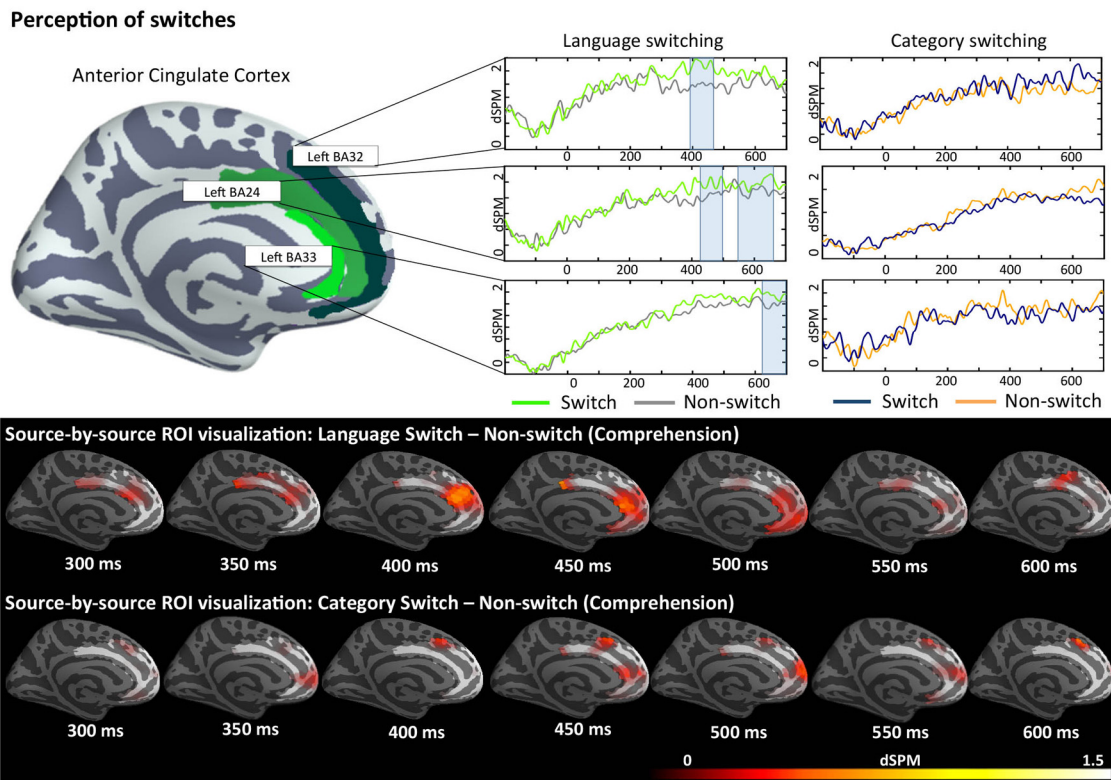
Figure 4. ROI results for pairwise comparisons for comprehension conditions in the left ACC activation averaged across subjects. A freeSurfer average brain illustrates the spatial distribution of the BAs included in the analyses color coded by BA. On the waveform plots, the shaded regions indicate that the difference in activity between the two tested conditions was significant at $p = 0.05$ (corrected), whereas the boxed region indicates marginally significant effects at $p < 0.01$ (corrected). Significance was determined using a non-parametric permutation test (Maris and Oostenveld, 2007) performed from 300–500 and 500–700 milliseconds (10,000 permutations). The whole-brain comparisons display all pairwise effects source by source within our ROIs. Only effects in which the first condition (switch condition) elicited greater activity than the second condition (non-switch condition) for at least 15 ms in at least 15 adjacent cortical sources at $p < 0.05$ are displayed, providing a complementary visualization of the spatial extent of our effects within the ROIs.

Post hoc analyses

To address the extent of the neural dissociation of switch effects between production and comprehension within language-switching, we ran 2 × 2 ANOVAs (response modality: production/comprehension and switch: switch/non-switch) in the dlPFC and ACC. The results of the analyses in the dlPFC revealed reliable interactions between response modality and switch in right BA10 (341–397 ms; $p = 0.04$) and right BA46 (422–460 ms; $p = 0.03$) and close to reliable in right BA9 (442–500 ms; $p = 0.09$), caused by increased activity for language switch over non-switch trials in production, but no difference between switch and non-switch trials in comprehension. The analyses of the left ACC revealed interaction clusters in left BA24 (588–668 ms; $p = 0.03$), left BA32 (402–428 ms; $p = 0.05$ and 626–659 ms; $p = 0.05$), and left BA33 (610–649 ms; $p = 0.05$) caused by increases in activity for switch over non-switch trials in comprehension, but no differences between conditions in production.

To investigate the extent to which the magnitude of the switch costs varied as a factor of the language in use, we ran a 2 × 2 ANOVA (language: Arabic/English and switch: switch/non-switch) within each domain (production and comprehension). Based on previous analyses, we targeted the ACC in comprehension and the dlPFC in production. The results of the analyses on language comprehension revealed a reliable interaction between language and switch in the left BA32 (404–454 ms; $p = 0.007$) caused by a greater increase for switch over non-switch trials in English (356–459 ms; $p = 0.02$) than in Arabic (334–352 ms; $p = 0.5$). An interaction cluster revealing a larger switch cost for Eng-

lish was also found in BA24 (434–455 ms; $p = 0.08$), but it did not reach reliability. The analyses on language production in the dlPFC revealed that the magnitude of the switch cost did not significantly diverge between Arabic and English in production (all interaction clusters in either direction, $p > 0.5$).

Additional ROI analyses

The 1 × 3 ANOVA within each of the tasks targeting possible gradual effects of switching did not reveal any reliable cluster of activation (all clusters $p > 0.1$).

Sensor data analyses

The sensor analysis indicated nonreliable increases in activity for switch over non-switch trials in the right frontal quadrant during language production and in the left front quadrant during domain-general production, resulting in waveform patterns qualitatively similar to the ones observed in the dlPFC ROI analysis. During language comprehension, nonreliable increases for switch conditions were observed in the left front and back quadrants, as well as in the right back quadrant. The *t*-map topography revealed that sensors containing increases for the switch condition clustered around the center part of the sensor array, thus accounting for the widespread distribution of the switch increases (Fig. 5). Finally, no increases were found in any of the quadrants during domain-general switching in comprehension. None of the mentioned increases were statistically significant; however, the results are plotted in Figure 5 for descriptive purposes.
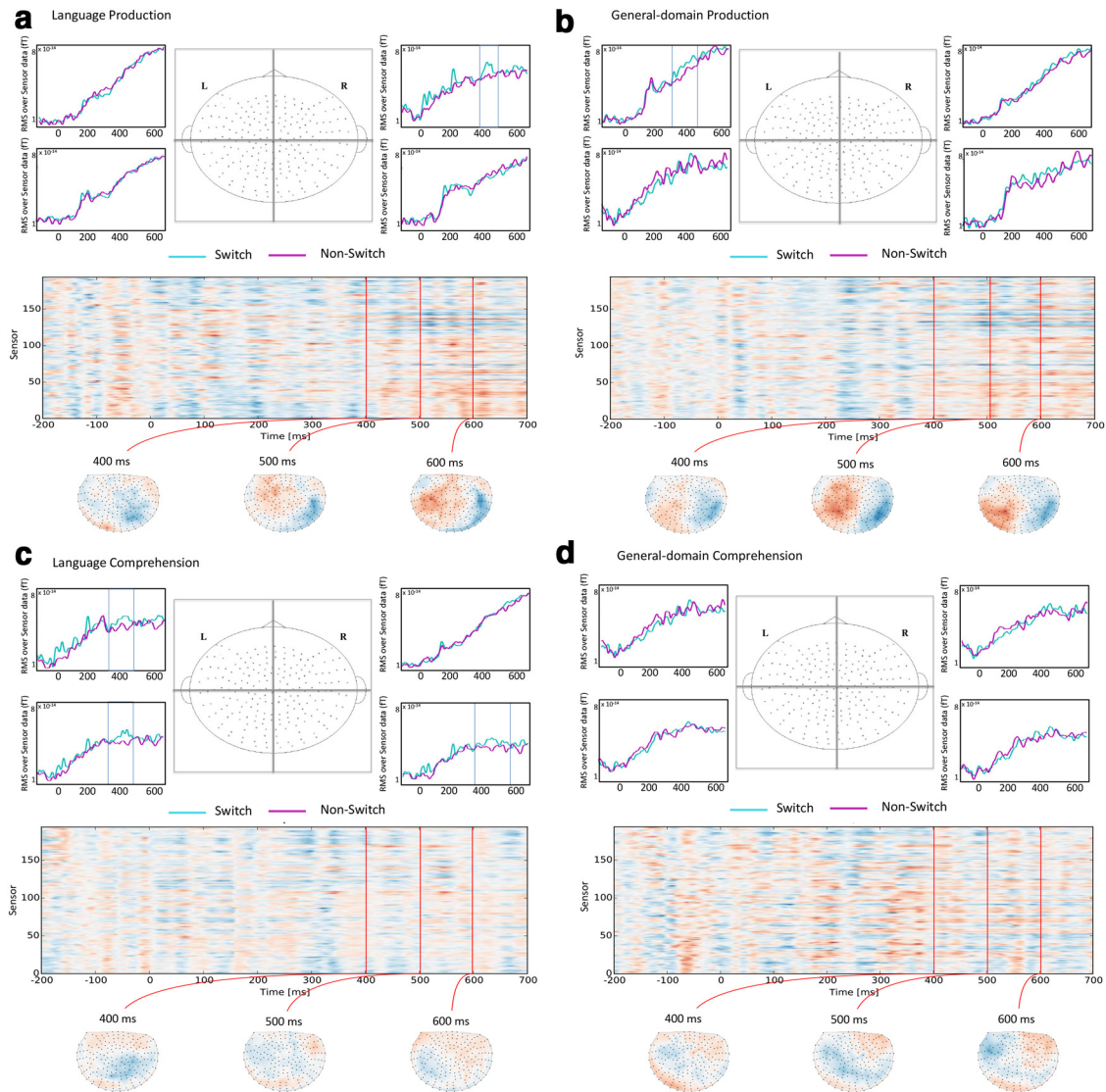
Figure 5. Pairwise comparisons of sensor data between switch and non-switch trials in all experimental tasks: (***a***) language production, (***b***) domain-general production, (***c***) language comprehension, and (***d***) domain-general comprehension. In each panel, the top part includes the activity waveforms for switch and non-switch conditions in each of the four partitions of the sensor array (front left, front right, back left, back right). A topography located in the middle of the four waveform plots shows the boundaries of the sensor array partitions. No pairwise comparison resulted in statistically reliable differences; however, clusters containing activity increases for switch conditions are displayed in a boxed region. The lower part of each panel includes a sensor/time distribution of *t*-values for the switch versus non-switch pairwise comparisons. Below this plot, the topographic distribution of such values is displayed: Red indicates increases for switch conditions and blue indicates increases for non-switch conditions.

Discussion

This work addressed the neurophysiological and behavioral effects of producing and perceiving language switches and investigated the anatomical overlap between language control mechanisms and high-level cognitive control more generally. Crucially, our experimental design directly contrasted closely parallel language and category-switching tasks performed in both production and comprehension, allowing for maximally informative comparisons.

Our results revealed a dissociation between language control mechanisms in production and comprehension, localizing the switch effect in language production in the dlPFC and in comprehension in the left ACC. Therefore, our findings on language-switching suggest that bilingual individuals do not rely on similar inhibitory mechanisms in production and comprehension; that is, there is no unique language "switch." Instead, our data suggest that the control networks facilitating production and compre-

hension processes diverge between at least auditory word recognition and word production. Therefore, it would appear that bilingual individuals rely on adaptive cognitive control in language-switching (Green and Abutalebi, 2013). Crucially, our results suggest that, in addition to the context in which an interaction occurs (Green and Abutalebi, 2013), the role a bilingual individual fulfills within that interaction also determines the neural networks involved, relying on the dlPFC when performing language-switching in production and on the ACC when perceiving switches in comprehension.

A possible explanation for the dissociation reported here is that, in comprehension, bilingual individuals' attention may focus on tuning the system to detect critical features that discriminate one language from another (Kuipers and Thierry, 2010; Krizman et al., 2012; Green and Abutalebi, 2013) and on monitoring relative language activation, whereas, in production, individuals may require a bias mechanism that enables the retrieval of

the target representation among other irrelevant candidates. If this is true, then switch costs in production would arise from the need to overcome the inhibition applied to the current target language in the preceding trial, whereas, in comprehension, switch costs would stem from the need to tune attention to detect the target language and subsequently strongly activate that language (Abutalebi et al., 2007). This interpretation predicts parallel switch costs for both languages in production for highly balanced bilingual individuals (Costa and Santesteban, 2004; Christoffels et al., 2007), but a larger switch cost when switching into the weaker language in comprehension (for a discussion, see Abutalebi et al., 2007). Our analyses of the switch-cost magnitude matched this prediction, consistent with previous research (Alvarez et al., 2003; Proverbio et al., 2004; Abutalebi et al., 2007) and supporting the hypothesis that only switching in production may rely on top-down inhibition, whereas switching in comprehension may instead require differential activation of each language (Dijkstra and Van Heuven, 1998, 2002; van Heuven et al., 2008). It follows from this that the ACC-based attention and monitoring network should be involved in comprehension (for review, see Carter et al., 1998, 1999; Botvinick et al., 2001; Braver et al., 2001; Ridderinkhof et al., 2004; Weissman et al., 2005), whereas production would require the intervention of the dlPFC, which is reportedly involved in response selection and inhibition (Abutalebi and Green, 2007). A compatible possibility is that, in the current paradigm, the dlPFC operates proactively in production (retrieving the target lexical item while maintaining the task goal indicated by the color cue) and the ACC reactively in comprehension (Hikosaka and Isoda, 2010). Along these lines, Luk et al. (2011) proposed that early activation in response to a cue may be sufficient to trigger proactive control in bilingual individuals. However, our results do not provide conclusive data in this respect and future research will have to test this specific hypothesis.

Importantly, the involvement of the dlPFC and the ACC replicates previous studies relating language-switching in production to the left (Hernandez et al., 2000; Rodriguez-Fornells et al., 2005; Wang et al., 2007) and right (Hernandez et al., 2001; Wang et al., 2007) dlPFC, and in comprehension to the left ACC (Abutalebi et al., 2007). The language switch effects localized in the dlPFC in production also extended to category-switching, consistent with previous studies (De Baene et al., 2015) and the hypothesis that language control is a subdomain of general cognitive control (Craik and Bialystok, 2006; Garbin et al., 2010; Abutalebi et al., 2012). Perhaps an unexpected aspect of our results is that there was a slight disparity in the extent to which each hemisphere was involved in each domain: whereas language control mostly recruited the right dlPFC, domain-general control recruited the left dlPFC. Although this differentiation falls within the usual range of observed effects (e.g., see Hernandez et al., 2001, Hosoda et al., 2012, and Wang et al., 2007 for right dlPFC language control effects; MacDonald et al., 2000, Perret, 1974 for left lateralized domain-general effects), there is also abundant evidence in the opposite direction, making this hemispherical distinction far from conclusive (e.g., see Hernandez et al., 2000, Rodriguez-Fornells et al., 2005, and Wang et al., 2007 for left dlPFC participation in language control and Egner and Hirsch, 2005 and Kuhl et al., 2007 for right dlPFC involvement in domain-general control). Moreover, the hypotheses regarding a possible distinction between the roles of the left and right dlPFC (Grafman et al., 2006; Kaller et al., 2011) are not straightforwardly applicable to the current case. Specifically, such hypotheses propose lateralization of prefrontal functions to be determined by their distinct involvement in the anticipatory

planning of actions. Given that, in our experiment, MEG activity was measured at the same stage of planning for both conditions, this proposition cannot accurately account for our results. Therefore, given the mixed results found in previous literature and the fact that in our study the main effect of switch was reliable across both hemispheres, we will refrain from making a distinction in this respect. Within the production modality, then, our findings support the basic assumption behind the bilingual advantage hypothesis: i.e., that there is a close relationship between language control and general cognitive control.

Conversely, effects of language-switching in comprehension in the left ACC did not extend to category-switching, suggesting a potential specialization of the ACC for language-switching (Branzi et al., 2015). However, these studies do not yet warrant strong conclusions given that there is in fact abundant prior evidence for the engagement of the ACC in non-language tasks (for review, see Carter and Van Veen, 2007), as well as shared involvement between language and domain-general cognitive control (De Baene et al., 2015). In our particular design, it could be the case that the temporal properties of the effects within category-switching might have precluded us from detecting shared switch-effects. In fact, Kieffaber and Hetrick (2005) found that category-switching effects in a similar paradigm occurred between 600 and 900 ms, an interval that unfortunately falls outside of the interval of analyzable data in the current study (participants were instructed to blink after the end of the auditory stimulus). Therefore, at this point, our findings simply raise the possibility of a language-specific ACC-switching effect in comprehension.

Our experimental design additionally allowed us to evaluate whether increased cognitive effort related to switching was constrained to the first element after a switch or if it carried over to following trials. Behavioral results showed that delayed naming latencies and increased error rates disappear gradually, as opposed to immediately, after a switch (Fig. 2), thus establishing for the first time the existence of a gradual switch effect in two parallel tasks, which provides support for the carryover account (Allport et al., 1994). However, this pattern was not observed in the MEG data, so further study will be needed to address the neurophysiological details of such an effect.

Last, the use of MEG allowed us to obtain the first characterization of the spatiotemporal profile of switching effects, establishing that such processes begin ~400 ms after stimulus presentation both in production and comprehension. The interval during which our effects were obtained was contained within the observed timing in prior ERP results on both domain-general rule switching (Karayanidis et al., 2003; Kieffaber and Hetrick, 2005) and language-switching (Jackson et al., 2001; Christoffels et al., 2007; Verhoef et al., 2009). Of particular interest is the fact that the switch effect in comprehension occurred later than the effects of language change identification (N2; 280–320 ms) in Kuipers and Thierry (2010), in which the time course of language change detection in English–Welsh bilingual individuals was investigated. However, it overlaps with their P600 (570–600 ms) effect, which they related to the reevaluation of the input to maintain the integrity of speech comprehension. Therefore, it is likely that the perceptual identification of phonological or prosodic features that identify auditory input as belonging to a different language occurs in an early time window (in the auditory cortex or superior temporal gyrus by hypothesis), but, critically, the 400–650 ms window indexes stimulus reevaluation (Osterhout and Holcomb, 1992; Hahne and Friederici, 1999) and plausibly top-down control to restrain lexical access to the appropriate lexicon. Our results confine spatially both of these processes to

the ACC. Therefore, the current study was able to locate the network of language control processes and to accompany its temporal profile with anatomical localizations.

Conclusion

This study demonstrates for the first time that the brain areas responsible for language-switching in production and comprehension dissociate even for identical lexical material. Although producing switches recruited the dlPFC bilaterally, comprehending them engaged the left ACC. This finding suggests that bilingual individuals rely on adaptive language control strategies and that the neural involvement during language-switching could be extensively influenced by whether the switch is performed or perceived. In addition, the observed anatomical overlap between language and category-switching in production suggests that language control is in fact a subdomain of general executive control in production (Craik and Bialystok, 2006; Garbin et al., 2010; Abutalebi et al., 2012), supporting the basic premise of the so-called bilingual advantage hypothesis (Bialystok et al., 2005) only in this domain.

Notes

Supplemental material for this article is available at http://www.psych.nyu.edu/pylkkanen/lab/estiblancoelorrieta_additional_materials.html, which contains participants' language background (adapted from Marian et al., 2007) and coordinates of all vertices in each BA label. This material has not been peer reviewed.

References

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

Abutalebi J, Brambati SM, Annoni JM, Moro A, Cappa SF, Perani D (2007) The neural cost of the auditory perception of language switches: an event-related functional magnetic resonance imaging study in bilingual individuals. J Neurosci 27:13762–13769. CrossRef Medline

Abutalebi J, Annoni JM, Zimine I, Pegna AJ, Seghier ML, Lee-Jahnke H, Lazeyras F, Cappa SF, Khateb A (2008) Language control and lexical competition in bilingual individuals: an event-related fMRI study. Cereb Cortex 18:1496–1505. CrossRef Medline

Abutalebi J, Della Rosa PA, Green DW, Hernandez M, Scifo P, Keim R, Cappa SF, Costa A (2012) Bilingualism tunes the anterior cingulate cortex for conflict monitoring. Cereb Cortex 22:2076–2086. Medline

Adachi Y, Shimogawara M, Higuchi M, Haruta Y, Ochiai M (2001) Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. IEEE Trans Appl Supercond 11:669–672. CrossRef

Allport A, Styles E, Hsieh S (1994) Shifting intentional set: exploring the dynamic control of tasks. In: Attention and performance XV: conscious and non-conscious information processing (Umiltà C, Moscovitch M, eds), pp 421–452. Cambridge, MA: MIT.

Alvarez RP, Holcomb PJ, Grainger J (2003) Accessing word meaning in two languages: an event-related brain potential study of beginning bilingual individuals. Brain Lang 87:290–304. CrossRef Medline

Bemis DK, Pylkkänen L (2011) Simple Composition: An MEG investigation into the comprehension of minimal linguistic phrases. J Neurosci 31:2801–2814. CrossRef Medline

Bemis DK, Pylkkänen L (2012) Combination across domains: An MEG investigation into the relationship between mathematical, pictorial, and linguistic processing. Front Psychol 3:583. Medline

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188. CrossRef

Bialystok E, Craik FI, Grady C, Chau W, Ishii R, Gunji A, Pantev C (2005) Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. Neuroimage 24:40–49. CrossRef Medline

Blanco-Elorrieta E, Pylkkänen L (2015) Brain bases of language selection: MEG evidence from Arabic–English bilingual language production. Front Hum Neurosci 9:27. Medline

Blanco-Elorrieta E, Pylkkänen L (2016) Composition of complex numbers:

delineating the computational role of the left anterior temporal lobe. Neuroimage 124:194–203. Medline

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychol Rev 108:624–652. CrossRef Medline

Branzi FM, Della Rosa PA, Canini M, Costa A, Abutalebi J (2015) Language control in bilingual individuals: monitoring and response selection. Cereb Cortex. In press.

Braver TS, Barch DM, Gray JR, Molfese DL, Snyder A (2001) Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. Cereb Cortex 11:825–836. CrossRef Medline

Calabria M, Hernández M, Branzi FM, Costa A (2011) Qualitative differences between bilingual language control and executive control: evidence from task-switching. Front Psychol 2:399. Medline

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280:747–749. CrossRef Medline

Carter CS, Botvinick MM, Cohen JD (1999) The contribution of the anterior cingulate cortex to executive processes in cognition. Rev Neurosci 10:49–57. Medline

Carter CS, van Veen V (2007) Anterior cingulate cortex and conflict detection: an update of theory and data. Cogn Affect Behav Neurosci 7:367–379. CrossRef Medline

Christoffels IK, Firk C, Schiller NO (2007) Bilingual language control: An event-related brain potential study. Brain Res 1147:192–208. CrossRef Medline

Costa A, Santesteban M (2004) Lexical access in bilingual speech production: evidence from language-switching in highly proficient bilinguals and L2 learners. J Mem Lang 50:491–511.

Craik FI, Bialystok E (2006) Cognition through the lifespan: mechanisms of change. Trends Cogn Sci 10:131–138. CrossRef Medline

Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, Aso T, Urayama S, Fukuyama H, Stockton K, Usui K, Green DW, Price CJ (2006) Language control in the bilingual brain. Science 312:1537–1540. CrossRef Medline

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26:55–67. CrossRef Medline

De Baene W, Duyck W, Brass M, Carreiras M (2015) Brain circuit for cognitive control is shared by task and language-switching. J Cogn Neurosci 27:1752–1765. CrossRef Medline

de Bruin A, Treccani B, Della Sala S (2015) Cognitive advantage in bilingualism an example of publication bias? Psychol Sci 26:99–107. CrossRef Medline

Del Prato P, Pylkkänen L (2014) MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. Front Psychol 5:524. Medline

Dijkstra T, Van Heuven WJ (1998) The BIA model and bilingual word recognition. In: Localist connectionist approaches to human cognition (Grainger J, Jacobs AM, eds), pp 189–225. London: Erbaum.

Dijkstra T, Van Heuven WJ (2002) The architecture of the bilingual word recognition system: From identification to decision. Bilingualism: Language and Cognition 5:175–197.

Egner T, Hirsch J (2005) Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. Nat Neurosci 8:1784–1790. CrossRef Medline

Garbin G, Sanjuan A, Forn C, Bustamante JC, Rodriguez-Pujadas A, Belloch V, Hernandez M, Costa A, Avila C (2010) Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. Neuroimage 53:1272–1278. CrossRef Medline

Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15:870–878. CrossRef Medline

Grafman J (2006) Planning and the brain. In: The human frontal lobes: functions and disorders (Miller BL, Cummings JL, eds), pp 249–261. London: Guildford.

Green DW, Abutalebi J (2013) Language control in bilingual individuals: the adaptive control hypothesis. J Cogn Psychol 25:515–530. CrossRef

Gross J, Baillet S, Barnes GR, Henson RN, Hillebrand A, Jensen O, Jerbi K, Litvak V, Maess B, Oostenveld R, Parkkonen L, Taylor JR, van Wassenhove V, Wibral M, Schoffelen JM (2013) Good practice for conducting

and reporting MEG research. Neuroimage 65:349–363. CrossRef Medline

Guo T, Liu H, Misra M, Kroll JF (2011) Local and global inhibition in bilingual word production: fMRI evidence from Chinese–English bilingual individuals. Neuroimage 56:2300–2309. CrossRef Medline

Hahne A, Friederici AD (1999) Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. J Cogn Neurosci 11:194–205. CrossRef Medline

Hämäläinen MS, Ilmoniemi RJ (1994) Interpreting magnetic fields of the brain: minimum norm estimates. Med Biol Eng Comput 32:35–42. CrossRef Medline

Hernandez AE, Martinez A, Kohnert K (2000) In search of the language switch: an fMRI study of picture naming in Spanish–English bilingual individuals. Brain Lang 73:421–431. CrossRef Medline

Hernandez AE, Dapretto M, Mazziotta J, Bookheimer S (2001) Language-switching and language representation in Spanish–English bilingual individuals: An fMRI study. Neuroimage 14:510–520. CrossRef Medline

Hikosaka O, Isoda M (2010) switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. Trends Cogn Sci 14:154–161. CrossRef Medline

Hosoda C, Hanakawa T, Nariai T, Ohno K, Honda M (2012) Neural mechanisms of language switch. J Neurolinguist 25:44–61.

Jackson GM, Swainson R, Cunnington R, Jackson SR (2001) ERP correlates of executive control during repeated language-switching. Bilingualism: Language and Cognition 4:169–178.

Kaller CP, Rahm B, Spreer J, Weiller C, Unterrainer JM (2011) Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. Cereb Cortex 21:307–317. CrossRef Medline

Karayanidis F, Coltheart M, Michie PT, Murphy K (2003) Electrophysiological correlates of anticipatory and postimulus components of task switching. Psychophysiology 40:329–348. CrossRef Medline

Kaufman EL, Lord MW, Reese TW, Volkmann J (1949) The discrimination of visual number. Am J Psychol 62:498–525. CrossRef Medline

Kieffaber PD, Hetrick WP (2005) Event-related potential correlates of task switching and switch costs. Psychophysiology 42:56–71. CrossRef Medline

Krizman J, Marian V, Shook A, Skoe E Kraus N (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci U S A 109:7877–7881. CrossRef Medline

Kuhl BA, Dudukovic NM, Kahn I, Wagner AD (2007) Decreased demands on cognitive control reveal the neural processing benefits of forgetting. Nat Neurosci 10:908–914. CrossRef Medline

Kuipers JR, Thierry G (2010) Event-related brain potentials reveal the time-course of language change detection in early bilingual individuals. Neuroimage 50:1633–1638. CrossRef Medline

Luk G, Green DW, Abutalebi J, Grady C (2011) Cognitive control for language-switching in bilingual individuals: A quantitative meta-analysis of functional neuroimaging studies. Language and Cognitive Processes 27:1479–1488. CrossRef Medline

MacDonald AW 3rd, Cohen JD, Stenger VA, Carter CS (2000) Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288:1835–1838. CrossRef Medline

Magezi DA, Khateb A, Mouthon M, Spierer L, Annoni JM (2012) Cognitive control of language production in bilingual individuals involves a partly independent process within the domain-general cognitive control network: Evidence from task-switching and electrical brain activity. Brain Lang 122:55–63. CrossRef Medline

Marian V, Blumenfeld HK, Kaushanskaya M (2007) The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilingual individuals and multilinguals. J Speech Lang Hear Res 50:940–967. CrossRef Medline

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods 164:177–190. CrossRef Medline

Meiran N, Chorev Z, Sapir A (2000) Component processes in task switching. Cogn Psychol 41:211–253. CrossRef Medline

Miller EK (2000) The prefrontal cortex and cognitive control. Nat Rev Neurosci 1:59–65. CrossRef Medline

Osterhout L, Holcomb PJ (1992) Event-related brain potentials elicited by syntactic anomaly. J Mem Lang 31:785–806. CrossRef

Perret E (1974) The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. Neuropsychologia 12:323–330. CrossRef Medline

Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD (1999) Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage 10:15–35. CrossRef Medline

Proverbio AM, Leoni G, Zani A (2004) Language-switching mechanisms in simultaneous interpreters: an ERP study. Neuropsychologia 42:1636–1656. CrossRef Medline

Pylkkänen L, Bemis DK, Blanco Elorrieta E (2014) Building phrases in language production: An MEG study of simple composition. Cognition 133:371–384. CrossRef Medline

Ridderinkhof KR, van den Wildenberg WP, Segalowitz SJ, Carter CS (2004) Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn 56:129–140. CrossRef Medline

Rodriguez-Fornells A, van der Lugt A, Rotte M, Britti B, Heinze HJ, Münte TF (2005) Second language interferes with word production in fluent bilingual individuals: brain potential and functional imaging evidence. J Cogn Neurosci 17:422–433. CrossRef Medline

Salthouse TA, Fristoe N, McGuthry KE, Hambrick DZ (1998) Relation of task switching to speed, age, and fluid intelligence. Psychol Aging 13:445–461. CrossRef Medline

Saltzman IJ, Garner WR (1948) Reaction time as a measure of span of attention. Am J Psychol 25:227–241.

Thompson-Schill SL, Swick D, Farah MJ, D'Esposito M, Kan IP, Knight RT (1998) Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. Proc Natl Acad Sci U S A 95:15855–15860. CrossRef Medline

van Heuven WJ, Schriefers H, Dijkstra T, Hagoort P (2008) Language conflict in the bilingual brain. Cereb Cortex 18:2706–2716. CrossRef Medline

van Heuven WJ, Dijkstra T (2010) Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. Brain Res Rev 64:104–122. CrossRef Medline

Verhoef K, Roelofs A, Chwilla DJ (2009) Role of inhibition in language-switching: evidence from event-related brain potentials in overt picture naming. Cognition 110:84–99. CrossRef Medline

Wagner AD, Koutstaal W, Maril A, Schacter DL, Buckner RL (2000) Task-specific repetition priming in left inferior prefrontal cortex. Cereb Cortex 10:1176–1184. CrossRef Medline

Wagner AD, Paré-Blagoev EJ, Clark J, Poldrack RA (2001) Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31:329–338. CrossRef Medline

Wang Y, Xue G, Chen C, Xue F, Dong Q (2007) Neural bases of asymmetric language-switching in second-language learners: an ER-fMRI study. Neuroimage 35:862–870. CrossRef Medline

Weissberger GH, Gollan TH, Bondi MW, Clark LR, Wierenga CE (2015) Language and task switching in the bilingual brain: Bilingual individuals are staying, not switching, experts. Neuropsychologia 66:193–203. CrossRef Medline

Weissman DH, Gopalakrishnan A, Hazlett CJ, Woldorff MG (2005) Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. Cereb Cortex 15:229–237. Medline

Westerlund M, Pylkkänen L (2014) The role of the left anterior temporal lobe in semantic composition vs semantic memory. Neuropsychologia 57:59–70. CrossRef Medline

Yuval-Greenberg S, Deouell LY (2009) The broadband-transient induced gamma-band response in scalp EEG reflects the execution of saccades. Brain Topogr 22:3–6. CrossRef Medline