

fMRI of Simultaneous Interpretation Reveals the Neural Basis of Extreme Language Control

Alexis Hervais-Adelman¹, Barbara Moser-Mercer², Christoph M. Michel³ and Narly Golestani¹

¹Brain and Language Lab, Department of Clinical Neuroscience, ²Faculty of Translation and Interpreting and ³Functional Brain Mapping Lab, Department of Fundamental Neuroscience, University of Geneva, Geneva, Switzerland

Address correspondence to Dr Alexis Hervais-Adelman, Brain and Language Lab, Campus Biotech, Neuroscience Department, 9 Chemin des Mines, Case Postale 60, Geneva 20, Switzerland. Email: alexis.adelman@unige.ch

We used functional magnetic resonance imaging (fMRI) to examine the neural basis of extreme multilingual language control in a group of 50 multilingual participants. Comparing brain responses arising during simultaneous interpretation (SI) with those arising during simultaneous repetition revealed activation of regions known to be involved in speech perception and production, alongside a network incorporating the caudate nucleus that is known to be implicated in domain-general cognitive control. The similarity between the networks underlying bilingual language control and general executive control supports the notion that the frequently reported bilingual advantage on executive tasks stems from the day-to-day demands of language control in the multilingual brain. We examined neural correlates of the management of simultaneity by correlating brain activity during interpretation with the duration of simultaneous speaking and hearing. This analysis showed significant modulation of the putamen by the duration of simultaneity. Our findings suggest that, during SI, the caudate nucleus is implicated in the overarching selection and control of the lexico-semantic system, while the putamen is implicated in ongoing control of language output. These findings provide the first clear dissociation of specific dorsal striatum structures in polyglot language control, roles that are consistent with previously described involvement of these regions in nonlinguistic executive control.

Keywords: caudate, cognitive control, language control, multilingualism, putamen

Introduction

Bilingualism confers a variety of cognitive advantages (Abutalebi et al. 2009; Diamond 2010), including improved non-linguistic executive skills (Bialystok et al. 2012), and delaying the appearance of symptoms of Alzheimer's disease (Schweizer et al. 2012). It has been suggested that the observed benefits of bilingualism stem from enhanced inhibitory control, required for continual selection between multiple languages in bilinguals (Green 1998). More recently, it has been proposed that the bilingual advantage arises from increased use of a more general "conflict monitoring" system in bilinguals compared with monolinguals (Costa et al. 2009; Hilchey and Klein 2011; Abutalebi et al. 2012b). Consistent with these suggestions, brain-imaging evidence indicates that bilingual language control depends upon a cortico-subcortical network incorporating the dorsal striatum, the anterior cingulate cortex (ACC) and the supplementary motor area (SMA) (Abutalebi and Green 2008; Kroll et al. 2008; Hervais-Adelman et al. 2011; Abutalebi et al. 2012a). These brain regions are known to participate in nonlinguistic executive-control processes such as response selection (Grahn et al. 2008), response inhibition

(Aron 2008), and conflict monitoring (Botvinick et al. 2004; Abutalebi et al. 2012b).

Electroencephalographic (EEG) research on language control has principally focused on the time course of event-related potentials, and has provided some results that are complementary to those of the functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) literature (Hervais-Adelman et al. 2011). The EEG literature is somewhat heterogeneous; it includes studies of perception and production with focuses on several different evoked components related to both language and executive function. This literature has been thoroughly reviewed and synthesized by Moreno et al. (2008). Among the domain-general components studied is the N200, which has been observed in several bilingual tasks (Rodriguez-Fornells et al. 2006); it is related to response suppression and has been localized to the ACC (Nieuwenhuis et al. 2003; Huster et al. 2010). Moreno et al. (2008) conclude that the event-related potential (ERP) evidence indicates that language switching requires active inhibition, and that the EEG signatures of language switching bear some similarity to those underlying withholding of responses in nonlinguistic paradigms.

Other ERP studies have allowed researchers to draw conclusions about the time course of language selection during speech production (for reviews see Costa 2005; Kroll et al. 2008). The weight of the evidence points toward simultaneous activation of a bilingual's two languages (e.g., Hoshino and Thierry 2011), in support of bilingual language models such as the bilingual interactive activation + model (Dijkstra and van Heuven 2002) that posit an integrated bilingual lexicon which is accessed in a language nonselective manner (van Heuven and Dijkstra 2010).

While great strides have been made in our understanding of polyglot language control, much of the existing evidence regarding the neural basis of bilingual language control comes from the study of tasks requiring only punctate applications of language control, such as language switching (Khateb et al. 2007; Wang et al. 2009; Garbin et al. 2011) and translation (Klein et al. 1995; Price et al. 1999; Quresima et al. 2002; Lehtonen et al. 2005). Distinct neural bases for general, executive task set maintenance, and for moment-to-moment cognitive control have previously been established (Dosenbach et al. 2008), and this distinction likely applies to polyglot language control. Given the mounting interest in the role of language control networks and the beneficial nonlinguistic consequences of their daily use, it is timely to explore the neural bases of extreme multilingual language control, using a task that places exceptional demands on both continuous and moment-to-moment control. We therefore examined the neural basis of SI in a group of 50 multilingual participants having had no previous SI experience. Successful execution of SI

depends heavily upon verbal working memory, simultaneous speech perception and articulation, and switching and dividing of attention between languages and between input and output modalities (Moser-Mercer et al. 2000). The existing neuroimaging literature on SI is very sparse; only 1 PET study on 8 professional interpreters revealed activation in the left inferior frontal gyrus and the SMA during interpretation (Rinne et al. 2000; Tammola et al. 2000).

Fabbro (1999) proposes a “neurofunctional” account of SI, which attempts to break down the task of SI into its cognitive subcomponents, and which incorporates a broad network of relatively left-lateralized brain areas. Fabbro proposes that acoustic analyses and motor output are handled bilaterally by the auditory and sensorimotor cortices, respectively. He suggests that phonological decoding takes place in left hemisphere subcortical and temporo-parietal structures, that semantic associations required to convert the message from one language to another occur in the left anterior inferior frontal cortex, and that articulatory processing for production is effected in left-lateralized subcortical structures, SMA and premotor regions, as well as in Broca’s area. The right hemisphere is characterized as being implicated in prosodic, emotional, and pragmatic decoding as well as in attentional control, for both production and perception. This account, however, focuses primarily on the mechanistic aspects of language perception and production implicated in SI, and does not directly concern itself with the question of the basis of language control or the executive components (such as working memory, planning, error monitoring, etc.) of the task. As outlined above, in addition to producing and comprehending speech, successful SI depends heavily on verbal working memory, simultaneous speech perception and production, and switching of attention between languages and between input and output modalities (Moser-Mercer 2000; Moser-Mercer et al. 2000).

Neuroimaging of SI provides a unique insight into the mechanisms of language control as it allows us both to examine the global network implicated by the task (task set maintenance), and to specifically test for brain regions whose activity is modulated by the duration of simultaneous speaking and listening (moment-to-moment control).

Materials and Methods

Participants

A group of 50 multilingual individuals (26 females, 7 left-handed, mean age: 25 years, range: 18–33 years) took part in the study. We present analyses of data from the 43 right-handed participants in the main text. For completeness, we additionally present analyses of data from the whole group, including the left-handed participants, in the Supplementary Materials. All participants had a high level of language proficiency in at least 3 languages. (Three languages are usually the minimum number required for admission to the Masters in Conference Interpretation offered by the Faculty of Translation and Interpretation at the University of Geneva. Early, balanced bilingual individuals are also admitted, if their proficiency in both their native languages is sufficiently high. Admission to the MA program is granted only upon successful passing of both written and oral proficiency examinations.) Twenty-three of the participants (of which 3 were left-handed) were enrolled in the Master’s degree in conference interpretation offered by the Faculty of Translation and Interpretation of the University of Geneva, and were scanned at the beginning of their training. (The first semester of the program focuses on consecutive interpretation. These participants were all scanned for the first time within a few weeks of

beginning their training, at least 3 months prior to beginning training in simultaneous interpretation.) The remaining participants were recruited from the student body of the University of Geneva. (Preliminary analysis of the data was performed in order to check for the presence of a main effect of group, or for a group by condition interaction. As expected, none was found, and the groups were combined for subsequent analyses.) Participants were asked to report the languages they speak or have studied, the age of acquisition of these, and to provide their own assessment of their fluency in each. A continuous language experience score was calculated for each individual, as follows: Knowledge of each language was weighted according to the age at which it had been learnt (higher weight for languages learnt earlier in life) and according to the fluency with which it was spoken (higher weight for more fluently spoken languages). The following weights were used: 1) proficiency: not fluent = 1, somewhat fluent = 2, moderately fluent = 3, quite fluent = 4, very fluent = 5, native = 6; 2) age of acquisition: ages ≥ 21 = 1, ages 13–20 = 2, ages 7–12 = 3, ages 1–6 = 4, at birth = 5 (this schema has previously been used by Golestani et al. 2011). All participants spoke English or French fluently, if not natively (see Table 1).

The experimental procedure was approved by the local research ethics committee (Geneva University Hospital, reference number: 09-161). Participants gave informed consent, and the experiment was conducted in accordance with the Declaration of Helsinki.

Design and Materials

Participants were scanned using a Siemens 3-T Trio MRI scanner with a 12-channel head coil. Stimuli were presented through MRI-compatible MR ConFon electrodynamic headphones. A sparse imaging procedure (Hall et al. 1999) was employed, enabling us to both sample the peak of the hemodynamic response function, and to present stimuli and record responses during silent intervals between image acquisitions.

In order to distinguish the neural basis of SI from that of simultaneous listening and speaking within one language (i.e., monolingual language control), we compared the brain activity elicited by simultaneous interpretation (SI) with that elicited by simultaneous repetition of speech (known as shadowing, SH). We also included a more basic, passive listening (PL) control condition.

In the “Listen” condition, participants were asked simply to listen attentively to the sentences being presented, and to make no response. In the “Shadow” condition, they were asked to overtly repeat the sentences they heard, beginning their response as soon as possible after the onset of each sentence (i.e., to speak while simultaneously hearing each sentence). In the “Interpret” condition, they were instructed to overtly simultaneously interpret the content of the sentences they heard into their most fluent language, which was almost always their native tongue. Again, participants were instructed to begin their responses as quickly as possible, and ideally before the input stream had stopped. Verbal responses were recorded for offline analyses aimed at ensuring that participants complied and that they accurately performed the interpretation task.

Professional simultaneous interpreters usually, though not invariably, work from a highly proficient language into their native tongue, which is considered to be substantially the easier direction (Fabbro 1999). We wished to ensure that the task was achievable, even for participants with no experience of SI. We therefore asked participants to interpret from the source language (i.e., that in which the stimuli were presented) into their native language (the target language). The source language was always either English or French, according to the participants’ preference. All conditions were carried out with the same source language for any given participant. This source language was always very well mastered (i.e., it was a very fluent non-native language), and for 2 early bilingual participants it was one of their 2 native languages [L1]. The following 9 L1s were represented: English, French, German, Spanish, Italian, Russian, Romanian, Portuguese, and Lithuanian (see Table 2 for a list of the source-target language combinations).

Stimuli consisted of a set of 156 French or English sentences. These were grammatically simple, having no more than one embedded clause, and grouped into quartets composing a four-sentence scenario, allowing participants to generate context-based expectation about the

Table 1

Language proficiency scores, list of languages, and number of languages

Subject	L1	L2	L3	N lang	L score	Source language	AoA of source	H
S1	RU	EN	FR	5	33	fr	7	R
S2	ES	EN	IT	7	49	en	7	R
S3	DE	IT	FR	6	54	fr	15	R
S4 ^c	NL ^a	FR ^a	EN	6	46	en	12	R
S5	EN	FR	ES	4	28	fr	14	R
S6	FR	ES	EN	5	39	en	11	R
S7	DE	EN	ES	5	38	en	8	R
S8	ES	FR	EN	6	38	fr	11	R
S9	IT	EN	DE	5	39	en	7	L
S10	RU	FR	EN	4	31	fr	12	R
S11	FR	DE	EN	7	51	en	9	R
S12	IT	EN	FR	5	37	en	6	R
S13	FR ^a	DE ^a	EN	4	34	fr	0	R
S14	IT	FR	EN	3	27	fr	6	R
S15	PT ^a	FR ^a	EN	3	30	fr	0	R
S16	ES	FR	EN	4	34	fr	10	L
S17	IT ^a	FR ^a	EN	3	30	fr	0	R
S18	ES	EN	SV	4	28	en	15	R
S19	IT ^a	FR ^a	PT	4	37	fr	0	R
S20	ES	FR	EN	4	33	fr	12	L
S21	IT	FR	EN	4	33	fr	7	R
S22	DE	FR	EN	9	59	fr	10	R
S23	ES	FR	EN	3	24	fr	16	R
S24	FR	DE	EN	7	48	en	11	R
S25	ES	EN	DE	5	35	en	10	R
S26	PT	FR	EN	4	37	fr	15	R
S27	FR ^a	EN ^a	ES	3	29	en	0	R
S28	IT	EN	DE	4	27	en	12	L
S29	DE	FR	EN	3	27	fr	10	R
S30	ES ^a	FR ^a	EN	3	28	fr	0	R
S31	LT	EN	DE	5	39	en	10	R
S32	RO	FR	IT	6	51	fr	4	L
S33	FR	IT	DE	4	34	en	16	R
S34	RU	EN	FR	4	33	fr	20	R
S35	RO	ES	FR	5	39	fr	8	R
S36	FR ^a	DE ^a	ES	4	37	fr	0	L
S37	ES	EN	FR	3	28	fr	16	R
S38	DE	FR	ES	4	31	fr	12	R
S39	ES	FR	DE	4	32	fr	16	R
S40	DE ^a	RM ^a	FR	6	55	fr	11	R
S41	IT	FR	EN	4	34	fr	8	R
S42	FR	ES	EN	7	51	en	11	R
S43	CAT ^a	ES ^a	EN	4	38	en	10	L
S44	FR	EN	IT	4	32	en	12	R
S45	IT	DE	FR	4	38	fr	12	R
S46	IT	FR	EN	4	33	fr	7	R
S47 ^b	HR	FR	RU	5	46	en	8	L
S48	IT	EN	FR	3	28	fr	11	R
S49	IT	FR	EN	4	31	en	11	R
S50	FR	EN	ES	3	27	en	12	R

Notes: Languages are designated using ISO-639 codes: EN, English, FR, French, IT, Italian, ES, Spanish, RU, Russian, DE, German, RO, Romanian, RM, Romansch, HR, Croatian, LT, Lithuanian, SV, Swedish, CAT, Catalan.

L1, L2, and L3 designate first, second, and third most fluent language, respectively. N lang is the total number of languages reported as somewhat proficient by the participants. L score is a language proficiency score, calculated on the basis of the participants' self-report as follows: Knowledge of each language was weighted according to the age at which it had been learnt (higher weight for languages learnt earlier in life) and according to the fluency with which it was spoken (higher weight for more fluently spoken languages). The following weights were used:

1) proficiency: not fluent = 1, somewhat fluent = 2, moderately fluent = 3, quite fluent = 4, very fluent = 5, native = 6; 2) age of acquisition: ages $\geq 21 = 1$, ages 13–20 = 2, ages 7–12 = 3, ages 1–6 = 4, at birth = 5. Source language is the language in which the task was carried out.

AoA of source is the age of acquisition of the source language, in years. H is handedness.

^aDual L1 for early (raised in a bilingual environment prior to schooling) bilinguals. The language into which these participants interpreted is marked in bold.

^bThis participant was the only one not to carry out simultaneous interpretation into their mother tongue, but rather into their L2. However, it must be noted that the proficiency of this individual in the L2 was sufficiently high for them to be granted a place to study simultaneous interpretation into this L2 at the Faculty of Translation and Interpretation in Geneva. This indicates a high-level mastery of the language.

^cThis participant elected to execute the interpretation task into French (one of their 2 L1s).

content of subsequent phrases. The thematic content of the French and English sets was highly overlapping, though nonidentical, due to matching constraints. (An example of an English quartet: “The cargo

Table 2

Distribution of language combinations for the task

Source language	Target language	Number of participants	Of which LH
English	French	11	1
English	Spanish	6	1
English	Italian	5	1
English	German	5	
		27 from English	
French	Italian	6	
French	Spanish	5	2
French	Russian	3	
French	German	3	1
French	Romanian	2	1
French	Lithuanian	1	
French	Portuguese	2	
French	English	1	
	(Total: 9 target languages)	23 from French	

ship has been found several weeks after disappearing at sea,” “The crew are all alive and well, and are currently answering questions,” “The few details that have emerged have left more questions than answers,” “There are rumors of secret cargoes and ransom demands”; and an example of a French quartet: “Il y a beaucoup de spéculation sur les circonstances de l’incident,” “Nous soupçonnons que le requin a pu confondre le surfeur avec un poisson blessé,” “Le jeune homme a été grièvement blessé dans l’attaque,” “Les médecins devront l’opérer toute la nuit pour sauver ses jambes” [in English: “There is a lot of speculation about the circumstances of the incident,” “We suspect that the shark mistook the surfer for an injured fish,” “The young man was seriously injured in the attack,” “Doctors will have to operate all night to save his legs”].) Sentences were recorded by a bilingual male speaker of French and of Southern British English. Participants heard sentences in all conditions exclusively in one of the 2 available languages (see Table 2). Sentence quartets were presented in random order, and were randomly allocated to conditions.

Immediately prior to the onset of each quartet, participants were presented with an on-screen cue for the duration of the preceding MRI acquisition (2.1 s), consisting of an instruction: “Listen,” “Shadow,” or “Interpret,” which they carried out for the subsequent 4 sentences. Conditions were pseudorandomly ordered such that the same instruction occurred for no more than 3 quartets in succession. The design is illustrated in Figure 1.

Prior to the experiment, participants underwent a brief practice session inside the scanner during which they were familiarized with the tasks and with the speaker’s voice and accent. The experimenters monitored performance during practice to ensure that participants were capable of carrying out the interpretation task in a simultaneous fashion. We specifically verified that participants began their responses while the stimuli were ongoing, and that they were capable of producing relatively fluent output in the target language during interpretation.

In order to provide an adequate control for brain activations associated with non-native language perception and comprehension, as well as within-language output monitoring, and division of attention between speaking and listening, we elected to use SH in the non-native language (which would be the source language during the SI trials). This is potentially somewhat different to L1 SH, it may, for example, induce co-activation of L1 (as would be suggested by the BIA+ model of Dijkstra and van Heuven 2002). SH in the non-native language (the source language for SI) thus allows us to control for any associated automatic co-activation of L1 induced when hearing the non-native language, since this likely also occurs during interpretation. It additionally provides the best available control for the perception, comprehension, monitoring, and switching components of a SI task from the same source language, enabling us to separate those from the additional cerebral activations elicited by the multilingual nature of the interpretation task.

fMRI Data Acquisition, Processing, and Analysis

A total of 195 functional volumes were acquired; divided over 3 runs that lasted ~10 min each (T_2^* -weighted echo-planar imaging (EPI),

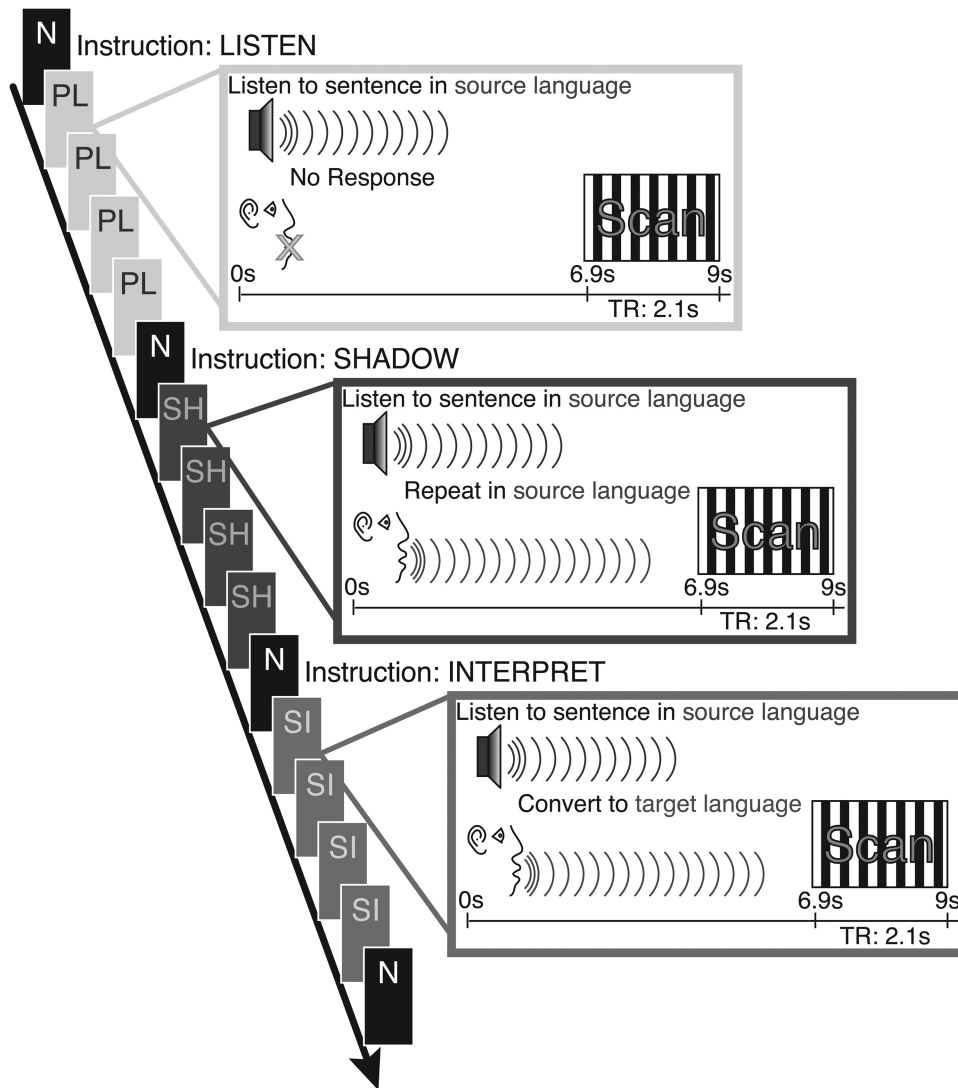


Figure 1. Illustration of fMRI paradigm. Main time-line represents sequence of events. N, null event, during which instructions were visually presented. PL, passive listening, SH, speech shadowing, SI, simultaneous interpretation. Panels illustrate composition of events. Every condition was presented in mini-blocks of 4 events (13 mini-blocks per condition, giving 52 scans per condition, plus 39 null events). The sentences presented within one mini-block were thematically linked and composed of a brief self-contained narrative. The order of conditions was pseudo-randomized such that the same condition never occurred for more than 3 mini-blocks in succession.

36 × 3.2 mm slices, 20% interslice gap, 3.2 × 3.2 mm in-plane resolution, angled away from the eyes to prevent ghost artifacts from aliasing of eye movements, TA: 2.1 s, TR: 9 s, TE: 80 ms). A total of 13 quartets were presented per condition over the course of the experiment, resulting in 52 trials per condition. Sentences were presented between scans, and every quartet was succeeded by a single null event in which no stimuli were presented, yielding 65 scans per run. Within each run, 13 different quartets of sentences were presented. Each run contained 4 occurrences of each of 2 conditions and 5 of the third. The allocation of these was constrained such that each of the conditions could occur 5 times in only 1, pseudorandomly assigned run, counterbalanced across participants. High-resolution (1 mm × 1 mm × 1.2 mm voxels) T_1 -weighted anatomical images were also acquired for each participant.

Preprocessing and analysis of the data were carried out in SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>). Data preprocessing for each subject included: 1) rigid realignment of each EPI volume to the first in the session, 2) co-registration of the structural volume to the mean EPI image, 3) normalization of the structural scan to a standard template (SPM8's single-subject T1 template) using the SPM8 unified segmentation routine, 4) normalization of all EPI volumes by application of the deformation parameters estimated in the normalization step, and 5) spatial smoothing using a Gaussian kernel of 8 mm.

First-level analysis was carried out using a general linear model for each participant in which every scan was coded for condition (Listen, Shadow, or Interpret), and null events were left unmodeled. For the 2 conditions in which participants made responses (shadow and interpret) the duration of overlap between listening and speaking for each trial was included as a parametric modulator for the condition in question. The duration of sentences was included as a regressor in the design matrix. Each run was modeled separately and the effect of block was coded separately. Each event was modeled using a single finite impulse response function in SPM8. Six parameters were appended in each run to code for the effects of movement (x , y , and z translations and x , y , and z rotations derived from the rigid realignment step of the preprocessing). A high-pass filter (cutoff 128 s) and AR1 correction for serial autocorrelation were applied. Contrast images for the pairwise comparisons of Shadow–Listen, Interpret–Shadow, and Interpret–Listen were produced for each participant, as were images showing the parametric modulation of activations during interpret and shadow trials by the duration of overlap.

Analysis of group data was achieved by entering contrasts of parameter estimates from single-subject models into random-effects analyses (single-sample t -tests), comparing differences of parameter estimates over subjects to zero. To account for potential differences due to

participants' different language expertise, their language experience scores were included as a covariate. The anatomical location of peaks was determined with reference to the automatic anatomical labeling (Eickhoff et al. 2009) and Brodmann area (BA) templates provided with the MRICron software package (<http://www.mccauslandcenter.sc.edu/mricro/mricron/>). All data reported reach a statistical significance threshold of whole-brain familywise error (FWE) corrected $P < 0.05$ at the voxel level.

Behavioral Responses

Participants' verbal responses were recorded and checked to ensure compliance with the instructions. Audio recordings were preprocessed to achieve noise-reduction and speech enhancement using Audacity (<http://audacity.sourceforge.net/>). Onsets and offsets of responses were determined using an automated detection procedure, achieved using an in-house Praat (<http://praat.org>, Boersma and Weenink 2011) script to scan wav files for increases in amplitude in combination with increases in zero-crossings of the waveform relative to baseline (similar to that implemented by Kello and Kawamoto (1998)). The duration of simultaneous speech listening and speaking (overlap) was calculated on the basis of the onsets and offsets of verbal responses and the known onsets and durations of the stimulus files.

Responses were scored by a panel of accredited professional simultaneous interpreters trained to work in the language combinations used during the interpretation and SH tasks by the participants (one rater per language). Raters were asked to evaluate SH responses on a binary scale to indicate failure to comply with instructions (score of 0) and compliance (score of 1). They were asked to assess the interpretation trials on a five-point scale as follows: 0 = no output, 1 = 1 content word, 2 = 2 content words (minimally a subject and object), 3 = 3 content words or more to make a meaningful interpretation, 4 = complete interpretation.

Results

Results presented here are for the 43 right-handed participants. For the sake of completeness, an analysis of all 50 participants, incorporating the 7 left-handed individuals, is provided in the Supplementary Materials.

Ratings of Performance

Due to equipment failure (microphone preamplifier glitches), only 46 complete sets of recordings out of 50 (including left-handed participants) were assessed. For the 43 right-handed participants included here, 41 complete datasets were evaluated. For these, interpretation and SH performance were highly satisfactory. Participants on average responded correctly and completely to over 90% of the SH trials (standard deviation [SD] over participants: 13.8), and attempted responses in over 95% of the interpretation trials (i.e., 5% missed trials or nonresponses). On average, participants' interpretations were rated at 3.16 (SD over participants: 0.49) on the scale described in the methods, indicating that average performance lay between completely accurate interpretations and highly meaningful responses. These results demonstrate that participants executed the task with a high degree of compliance and a good level of accuracy.

Neural Basis of SH: Monolingual Language Control

Comparing the activation elicited by SH with that elicited by PL allowed us to examine the neural systems that support monolingual language control arising from speech production and self-monitoring while simultaneously attending to an exogenous speech stream. The results of this contrast are shown in Table 3 and Figure 2, and are broadly concordant with

previously reported data on SH (Peschke et al. 2009) and on speech production (Peeva et al. 2010). We note that despite the SH having been in a language other than the participants' L1, the activation pattern does not seem qualitatively different from that previously reported for SH tasks. Further, we observe no significantly elevated activity in the caudate nuclei or in the putamen, subcortical structures that have previously been associated with L2 repetition tasks (Klein et al. 1994, 2006). For the sake of brevity, we will not further discuss the comparison of SH and listening in this article.

Neural Basis of SI: Bilingual Language Control

An initial comparison of activation during SI with that during PL revealed that SI recruits all the brain regions implicated in SH. Brain areas recruited by SI over and above those recruited by SH are involved in handling the additional demands of interpretation, which can be superficially characterized as the real-time conversion of the input stream from one language to another. Achieving this successfully depends upon multiple cognitive processes: deep but rapid semantic and syntactic analysis of the input, retrieval of appropriate lexical, syntactic and stylistic alternatives in the other language (i.e., translation), planning of the corresponding linguistic and semantic output, execution of speech output (speech production), monitoring of the output, error correction, and, finally, continuous attention to the incoming speech stream (Moser 1978). A comparison of responses observed during SI with those observed during SH revealed involvement of the left anterior SMA and pre-SMA, the left anterior insula, the left premotor cortex, the caudate nuclei, crus I of the right cerebellum and the dorsal ACC (dACC) [Fig. 2, Table 4]. Further significant increases in activation were also found in the left pars triangularis (BA 45, the anterior portion of Broca's area) and pars orbitalis (BA 47) of the inferior frontal gyrus.

Neural Correlates of Simultaneity

We wanted to not only examine the network that is engaged by the overall task requirements, but also to examine networks involved in the execution of moment-to-moment control during task performance. The above subtractive analyses summarize the brain activity associated with the relative demands of the respective experimental tasks, including maintaining and controlling the appropriate language set (a prerequisite of carrying out the task correctly), and the mechanisms implicated in effecting the task. In order to better understand the neural bases underlying moment-to-moment linguistic and cognitive control required during SI and also during SH, we performed analyses complementary to the above subtractive ones by testing for brain regions whose activation was modulated by the duration of simultaneous speaking and listening. During these periods of overlap, participants must concurrently process input and produce output, while monitoring the output for errors. The brain regions whose response is modulated by the duration of these periods of simultaneity are likely to be those that are implicated in the moment-to-moment control demanded by either SH or interpreting. The mean duration of overlap was 2.1 s (SD = 1.04 s) during SI, and 2.46 s (SD = 0.98 s) during SH.

During SH, the brain areas modulated by simultaneity are principally the superior temporal gyri, and a portion of the left SMA (see Fig. 3 and Table 5). A broader network is modulated

Table 3
Shadowing versus passive listening

Region name	Hem	MNI coordinates (mm)			T-score	z-Score	P_{FWE}	Cluster size (voxels)
		x	y	z				
Postcentral gyrus (BA3)	Left	-46	-12	36	15.45	Inf	<0.001	10868
Precentral gyrus (BA4)	Left	-56	0	22	13.35	Inf	<0.001	
Inferior frontal gyrus (pars opercularis) (BA44)	Left	-46	10	4	10.35	7.22	<0.001	
Anterior insula	Left	-38	8	4	10.2	7.16	<0.001	
Subthalamic nucleus	Left	-12	-18	4	10.11	7.12	<0.001	
Posterior superior temporal gyrus (BA42)	Left	-56	-40	18	9.81	7	<0.001	
Inferior frontal gyrus (pars opercularis)	Left	-54	14	8	9.58	6.9	<0.001	
Mid superior temporal gyrus (BA22)	Left	-48	-30	6	9.4	6.83	<0.001	
Superior temporal pole (BA38)	Left	-52	12	-6	9.34	6.8	<0.001	
Globus pallidus	Left	-22	-6	-4	9.11	6.7	<0.001	
Thalamus	Right	12	-18	0	9.07	6.68	<0.001	
Posterior superior temporal gyrus (BA22)	Left	-56	-26	6	8.97	6.64	<0.001	
Inferior frontal gyrus (pars opercularis) (BA44)	Left	-48	12	16	8.75	6.53	<0.001	
Planum temporale (BA41)	Left	-40	-32	10	8.56	6.45	<0.001	
Anterior insula	Left	-30	22	6	8.17	6.26	<0.001	
Globus pallidus	Right	16	-2	-4	8.01	6.18	<0.001	
Globus pallidus	Right	22	-8	-2	7.98	6.17	<0.001	
Brainstem white matter	Left	-8	-22	-10	7.94	6.15	<0.001	
Putamen	Left	-18	2	10	7.94	6.14	<0.001	
Superior temporal sulcus (BA22)	Left	-58	-40	8	7.93	6.14	<0.001	
Central sulcus (BA4)	Right	52	-8	26	15.37	Inf	<0.001	5011
Central sulcus (BA4)	Right	50	-10	40	13.95	Inf	<0.001	
Anterior insula	Right	40	12	2	10.12	7.12	<0.001	
Middle temporal gyrus	Right	52	-24	6	9.85	7.01	<0.001	
Middle temporal gyrus (BA22)	Right	64	-28	12	8.26	6.3	<0.001	
White matter subjacent to middle temporal gyrus	Right	44	-26	-10	6.67	5.46	0.001	
Superior temporal pole (BA38)	Right	54	14	-16	5.46	4.71	0.027	
Supplementary motor area (BA6)	Left	-6	0	64	15.23	Inf	<0.001	3076
Anterior cingulate cortex (BA24)	Left	-6	10	42	9.83	7.01	<0.001	
Anterior cingulate cortex (BA32)	Right	10	12	38	9.52	6.88	<0.001	
Cerebellum lobule VI	Right	16	-62	-20	12.25	Inf	<0.001	2692
Cerebellum lobule VI	Left	-14	-62	-20	11.55	7.66	<0.001	
Vermis VI	Left	0	-68	-14	8.81	6.56	<0.001	
Cerebellum lobule VI	Right	36	-54	-28	8.2	6.27	<0.001	
Central sulcus (between BA3 and BA4)	Left	-18	-30	60	7.4	5.87	<0.001	64
Calcarine sulcus (BA17)	Left	-10	-74	12	6.58	5.41	0.001	353
Cerebellum lobule VIII	Right	14	-66	-40	6.48	5.35	0.002	39
Cerebellum crus I	Left	-48	-60	-28	6.45	5.33	0.002	25
White matter [fronto-opercular fascicle]	Right	28	-34	10	5.93	5.01	0.007	108
White matter [corticospinal tract]	Right	22	-28	18	5.39	4.66	0.033	
Calcarine sulcus (BA17)	Right	14	-68	12	5.6	4.8	0.018	23
Precentral gyrus (BA3)	Right	20	-30	62	5.58	4.79	0.02	10

Table lists maximum of 20 voxel-peaks per cluster, spaced a minimum of 8 mm apart. Bold rows indicate peak voxel within cluster.
Hem, hemisphere; P_{FWE} , familywise error-corrected P -value.

by simultaneity during SI (see Fig. 3 and Table 6). In addition to the superior temporal gyri, it includes 2 left frontal regions (medial prefrontal and medial orbitofrontal cortices [mPFC and mOFC]), alongside the putamen and the superior aspect of the cerebellum bilaterally.

Dissociating the Roles of the Caudate and the Putamen

In order to establish whether there was a dissociation between the roles of the caudate and the putamen during SH and interpretation, an additional analysis was carried out. We selected the voxels in the left and right caudate and putamen showing the peak effect over the group for the contrast of interpretation versus SH and for modulation by simultaneity during interpretation. For each of the structures, we defined a region of interest in the left and right caudate as the cluster showing significant ($P_{FWE} < 0.05$) activation for interpretation-SH, and for the left and right putamen as the region showing significant modulation by simultaneity during interpretation. The first eigenvariate of the T-statistic for activation in these regions was extracted, subjectwise, for the tests: SH-baseline, interpreting-baseline, modulation by SH, and modulation by interpreting,

using Marsbar (Brett et al. 2002). These were then analyzed using a 3-way repeated-measures ANOVA with the following within-subject factors: structure (4 levels: left and right caudate and putamen), condition (2 levels: s SH and interpretation) and analysis type (2 levels: subtractive and modulation). Language experience was included as a covariate of no interest. A significant 3-way interaction was observed (degrees of freedom Greenhouse-Geisser corrected for nonsphericity $F_{1,382,56.049} = 4.267$, $P = 0.032$, partial $\eta^2 = 0.094$), indicating a difference in response in the different structures as a function of condition and analysis type (see Fig. 4). Post hoc comparisons show that in the right putamen, there was equal response to interpretation and SH versus baseline, and significantly greater modulation by the duration of overlap during interpretation than during SH (post hoc pairwise comparison: $T_{(42)} = 2.662$, $P = 0.011$). Similarly, in the left putamen, there was equal response to interpretation and SH versus baseline, and there was a marginally significant greater effect of simultaneity on modulation during interpretation than during SH (post hoc pairwise comparison: $T_{(42)} = 1.974$, $P = 0.055$). (We note that this difference is significant if we apply a directional hypothesis of “modulation by simultaneity is greater during

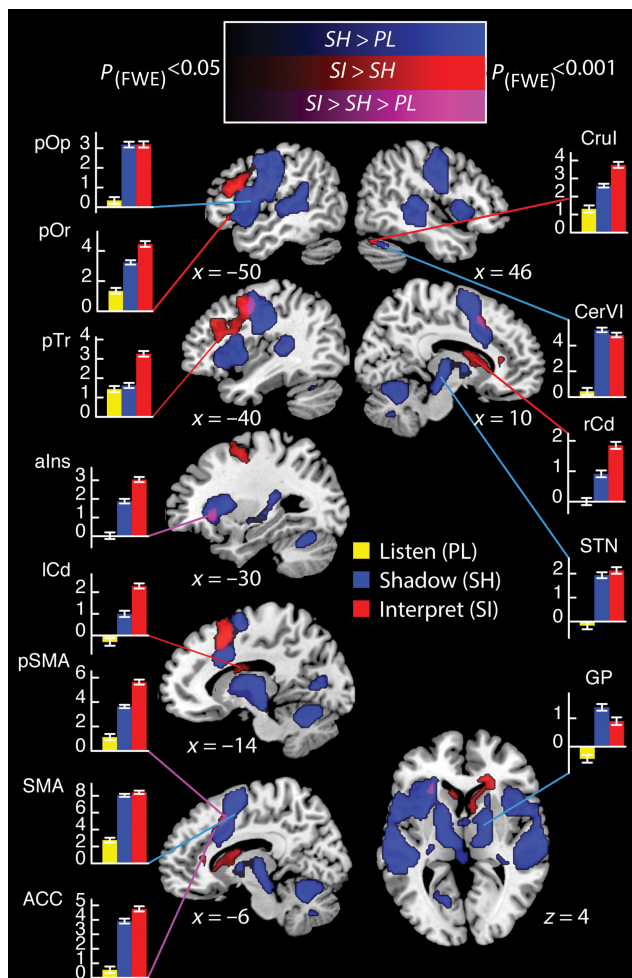


Figure 2. Significant differences in BOLD response for the contrasts shadowing > passive listening (blue) and simultaneous interpreting > shadowing (red). Regions where both contrasts are significant (i.e., simultaneous interpreting > shadowing > passive listening) are also shown (magenta). Results are projected on a canonical single-subject MNI brain. Shadowing induces activation in a broad bilateral network encompassing temporal, inferior frontal, inferior parietal, motor, and subcortical regions including thalamus and globus pallidus. Interpretation additionally recruits further inferior frontal, motor, and basal ganglia regions. Bar charts indicate T-statistics averaged over the group at selected peak voxels in the named activation clusters, and error bars indicate standard error of the mean corrected to be appropriate for repeated-measures designs (Loftus and Masson 1994). Coordinates indicate plane of section (in MNI space) of the background images. pOp, pars opercularis; pOr, pars orbitalis; pTr, pars triangularis; alns, anterior insula; ICd, left caudate; pSMA, presupplementary motor area; SMA, supplementary motor area; ACC, anterior cingulate cortex; Cru1, Crus 1 of cerebellum; CerVI, Cerebellum lobule VI; rCd, right caudate; STN, subthalamic nucleus; GP, globus pallidus.

interpretation than during shadowing” [i.e. using a one-tailed test].)

Discussion

Simultaneous Interpretation

All the regions involved in SH are also involved in SI, reflecting the common linguistic and executive demands of the tasks. Left inferior frontal gyrus regions additionally recruited during interpretation over SH include pars triangularis, known for its role in semantic processing (Dapretto and Bookheimer 1999; Bookheimer 2002), and pars orbitalis, implicated in semantic

memory and cognitive control of memory (Badre and Wagner 2007). Pre-SMA and dACC were also additionally recruited during SI. Pre-SMA activation has been associated with tasks in which there is response conflict, and generation of complex motor acts (Picard and Strick 1996; Nachev et al. 2008). Pre-SMA and dACC are thought to play complementary roles in action selection: The pre-SMA is involved in the selection and initiation of action sets, and the dACC in monitoring the outcome for errors (Rushworth et al. 2004). During SI, there is a high level of competition between languages, and the engagement of the pre-SMA is consistent with the need to handle the resulting response conflict. This analysis of the results is supported by previous studies on language switching, which have implicated the pre-SMA (Abutalebi et al. 2008).

The caudate nuclei were also engaged specifically during SI. The left caudate nucleus has previously been reported to be recruited in multilingual paradigms (Crinion et al. 2006; Abutalebi and Green 2008; Garbin et al. 2010), and has been shown to be implicated in multilingual language control by direct electrical stimulation (Wang et al. 2012).

The left anterior insula and SMA have been associated with a speech preparatory loop (Riecker et al. 2005). SI calls on speech preparation to a greater extent than SH due to the requirement to formulate the output *ab initio* rather than simply repeating the heard speech segments, and due to more effortful speech output due to the competition between 2 languages during the former, but not the latter, condition. The increased anterior insula and dACC activation may also reflect the increased attentional demands and difficulty of SI compared with SH, managed by the cingulo-opercular system (Petersen and Posner 2012).

The network of regions that we identified during SI is broader than previously reported. This could be due to our larger sample size, or may be due to the fact that participants in the present study were untrained, whereas the previously studied participants were professional simultaneous interpreters with between 5 and 20 years of experience (Rinne et al. 2000; Tommola et al. 2000). Trained interpreters’ expertise may enable them to carry out the task efficiently and effectively while recruiting fewer brain regions than naïve participants, consistent with many studies showing decreased brain activation when a task is more rehearsed and automated compared with when it is more effortful and novel (Ericsson et al. 2006).

The regions we identify include many of those proposed by Fabbro (1999) to underlie SI. In addition to these, our investigation reveals the involvement of brain regions involved in the language- and executive-control requirements of the task.

The Neural Correlates of Simultaneity

An important distinction exists between the cognitive control required to meet the overall demands of specific task instructions, and the exercise of control on a moment-to-moment basis while carrying out that task. This distinction is borne out in the existing neuroimaging literature, which shows different brain networks associated with task set maintenance over the full length of experimental trials, or associated with the initiation of responses and the adjustment of control on a moment-to-moment basis (Dosenbach et al. 2008). Our parametric modulation analysis aimed to reveal the latter. During periods of overlapping speech input and speech production, the participants are concurrently engaged in processing the

Table 4
Interpretation versus shadowing

Region name	Hem	MNI coordinates (mm)			T-score	z-Score	P_{FWE}	Cluster size (voxels)
		x	y	z				
Presupplementary motor area (BA6)	Left	-6	14	60	10.94	7.44	<0.001	2833
Premotor cortex (BA6)	Left	-40	2	48	7.65	6	<0.001	
Inferior frontal gyrus (pars triangularis)	Left	-48	20	24	7.56	5.95	<0.001	
Inferior frontal gyrus (pars opercularis) (BA44)	Left	-46	12	30	7.08	5.69	<0.001	
Superior frontal sulcus (BA6)	Left	-26	-2	54	6.46	5.34	0.002	
Inferior frontal sulcus (BA45)	Left	-42	32	32	6.07	5.1	0.005	
Anterior cingulate cortex (BA32)	Right	14	24	38	5.84	4.95	0.009	
Anterior cingulate cortex (BA32)	Right	14	12	48	5.84	4.95	0.009	
Superior frontal sulcus (between BA6 and BA8)	Left	-30	6	64	5.48	4.72	0.024	
Anterior insula (BA47)	Left	-28	26	-2	7.62	5.98	<0.001	86
White matter [inferior fronto-occipital fasciculus] (BA47)	Right	22	34	-2	6.73	5.49	0.001	375
White matter [anterior thalamic radiation] (BA47)	Right	22	28	6	6.73	5.49	0.001	
Head of the caudate	Right	12	20	6	5.99	5.05	0.006	
Head of the caudate	Right	10	4	16	5.93	5.01	0.007	
Head of the caudate	Right	8	12	8	5.92	5.01	0.007	
White matter [callosal body]	Right	12	32	8	5.85	4.96	0.009	
White matter [superior corona radiata]	Left	-20	4	30	6.03	5.07	0.005	72
Head of the caudate	Left	-14	0	22	5.94	5.02	0.007	
Cerebellum crus I	Right	48	-66	-28	5.87	4.97	0.008	12
Head of the caudate	Left	-8	20	4	5.83	4.95	0.009	19
Superior frontal sulcus	Right	24	0	50	5.6	4.8	0.017	14
White matter [superior corona radiata]	Right	22	4	28	5.59	4.79	0.018	22
Inferior frontal gyrus (between pars orbitals and pars triangularis)	Left	-52	22	-4	5.56	4.77	0.019	11

Table lists maximum of 20 voxel-peaks per cluster, spaced a minimum of 8 mm apart. Bold rows indicate peak voxel within cluster. Hem, hemisphere; P_{FWE} , familywise error-corrected P -value.

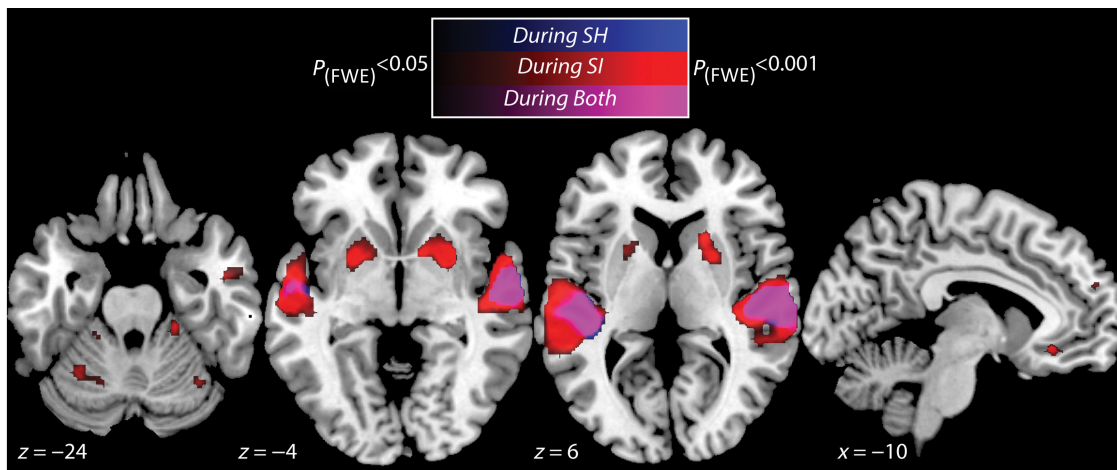


Figure 3. Regions showing significant modulation of BOLD response as a function of the duration of overlapping speaking and listening during shadowing (blue) and interpretation (red) and both (magenta), projected on a the canonical single-subject MNI brain. During these periods of overlap, participants are concurrently engaged in processing input, producing an output, and monitoring that output for errors. This modulation analysis thus reflects the neural bases of dynamic, moment-to-moment cognitive control required within each task. During both shadowing and interpretation, there is significant modulation of superior temporal lobe activity by simultaneity. During shadowing, a portion of the SMA also displays significant modulation. During interpretation, a broader network of regions, constituting a putamen-cerebellar-prefrontal circuit is modulated by simultaneity. Coordinates indicate plane of section (in MNI space) of the background images.

Table 5
Regions significantly modulated by simultaneity during shadowing

Region name	Hem	MNI coordinates (mm)			T-score	z-Score	P_{FWE}	Cluster size (voxels)
		x	y	z				
Planum temporale (BA41)	Left	-40	-30	12	9.82	7	<0.001	1090
Superior temporal sulcus (BA22)	Left	-60	-12	-2	6.93	5.61	0.001	
Superior temporal sulcus (between BA20 and BA22)	Left	-52	-10	-6	6.18	5.16	0.005	
Heschl's gyrus	Right	48	-18	2	9.57	6.9	<0.001	1339
Anterior superior temporal gyrus (BA21)	Right	64	-8	-2	9.53	6.88	<0.001	
Mid superior temporal gyrus (BA22)	Right	62	-20	4	8.85	6.58	<0.001	
Mid superior temporal gyrus	Right	56	-12	-2	8.58	6.46	<0.001	
Heschl's gyrus	Right	44	-22	10	8.19	6.27	<0.001	
Planum temporale (BA41)	Right	48	-30	12	6.84	5.56	0.001	

Table lists maximum of 20 voxel-peaks per cluster, spaced a minimum of 8 mm apart. Bold rows indicate peak voxel within cluster. Hem, hemisphere; P_{FWE} , familywise error-corrected P -value.

Table 6

Regions showing significant modulation of activity by simultaneity during interpretation

Region name	Hem	MNI coordinates (mm)			T-score	z-Score	P_{FWE}	Cluster size (voxels)
		x	y	z				
Anterior superior temporal sulcus (BA22)	Right	60	-4	-8	11.48	7.63	<0.001	3215
Heschl's gyrus	Right	44	-20	10	10.89	7.42	<0.001	
Mid superior temporal gyrus (BA22)	Right	52	-18	-2	9.44	6.84	<0.001	
Planum temporale (BA21)	Right	68	-22	2	9.04	6.67	<0.001	
Mid superior temporal gyrus (BA22)	Right	56	-22	10	8.66	6.49	<0.001	
Posterior superior temporal gyrus (BA22)	Right	64	-32	8	8.18	6.26	<0.001	
Posterior superior temporal sulcus (BA21)	Right	50	-38	4	7.15	5.73	<0.001	
Deep sylvian fissure [between parietal operculum and planum temporale] (BA48)	Right	34	-28	22	6.26	5.22	0.004	
Mid superior temporal sulcus (BA22)	Left	-52	-30	4	9.93	7.05	<0.001	
Anterior superior temporal sulcus (BA22)	Left	-58	-14	-4	9.8	6.99	<0.001	
Posterior superior temporal gyrus (BA22)	Left	-56	-36	8	8.16	6.25	<0.001	
Heschl's gyrus	Left	-48	-18	6	7.59	5.97	<0.001	
Anterior superior temporal sulcus (BA38)	Left	-52	4	-8	7.2	5.76	<0.001	
Planum temporale (BA42)	Left	-58	-34	16	6.44	5.32	0.002	
Deep sylvian fissure [between parietal operculum and planum temporale]	Left	-34	-28	18	6.01	5.06	0.008	
Anterior middle temporal gyrus (BA21)	Left	-52	-2	-18	5.39	4.66	0.043	
Putamen	Right	28	6	-4	8.29	6.32	<0.001	
White matter [anterior limb of internal capsule]	Right	20	16	6	7.3	5.81	<0.001	
White matter [anterior limb of internal capsule] (BA0)	Right	12	8	-4	7.21	5.76	<0.001	
Anterior ventral paracingulate cortex (BA11)	Left	-8	32	-10	7.6	5.97	<0.001	
Putamen	Left	-24	4	-4	7.33	5.83	<0.001	
Putamen	Left	-22	10	2	6.82	5.55	0.001	
White matter [anterior limb of internal capsule]	Left	-18	4	14	6.2	5.18	0.005	
Putamen (BA0)	Left	-24	2	8	5.62	4.81	0.024	
Cerebellum lobule IV/V (BA30)	Right	22	-32	-24	6.65	5.45	0.001	
Precentral gyrus (BA6)	Right	30	-16	50	6.34	5.26	0.003	
Cerebellum lobule VI (BA37)	Left	-28	-54	-24	6.31	5.24	0.003	
Cerebellum lobule VI (BA19)	Left	-16	-62	-24	5.77	4.91	0.015	
Medial superior frontal gyrus (BA10)	Left	-2	60	26	6.21	5.18	0.005	
Dorsal anterior cingulate gyrus (BA32)	Left	-10	54	24	6.07	5.1	0.007	
Cerebellum lobule VI	Right	36	-62	-26	6.14	5.14	0.005	
Cerebellum crus II	Right	10	-82	-38	6.11	5.12	0.006	
Cerebellum crus II	Right	22	-78	-38	5.6	4.8	0.025	
Superior parietal lobule (BA5)	Left	-16	-44	64	6.09	5.11	0.006	
Cerebellum lobule IV/V	Left	-20	-36	-22	5.92	5.01	0.01	
Inferior occipital cortex (BA18)	Left	-22	-94	-10	5.76	4.9	0.016	

Table lists maximum of 20 voxel-peaks per cluster, spaced a minimum of 8 mm apart. Bold rows indicate peak voxel within cluster.

Hem: hemisphere, P_{FWE} : familywise error-corrected P -value.

input stream, producing an output, and monitoring that output for semantic, syntactic and stylistic compatibility with the source language input. By revealing brain areas that are more heavily recruited during longer lasting periods of simultaneous task execution, the modulation analysis reflects the neural bases of dynamic, moment-to-moment cognitive control required within each task. These very regions are likely to be those that meet the greater linguistic and attentional control requirements of the tasks during the more demanding periods of simultaneity.

Simultaneity in both SH and SI modulated auditory regions (the superior temporal gyri). Greater overlap of the input and output speech streams results in a more complex auditory input for a longer period of time. Consequently, there are likely to be increased demands on auditory processing and on auditory attention during periods of overlapping speech perception and production.

During SI, we also found modulation of the left mPFC, which has been implicated in multitasking. It has been suggested that this region serves to balance attention between stimulus-oriented and stimulus-independent processes (Gilbert et al. 2006), a role likely to be called upon more when handling demands of processing the input stream, producing speech in the target language, and monitoring output simultaneously as opposed to consecutively. The mOFC also showed modulation by simultaneity and it may be involved in

language switching (Wang et al. 2007, 2009), another key component of interpretation.

During SI, activity in the putamen, bilaterally, was modulated by simultaneity, with the response of the right putamen showing a significant dissociation such that its response was significantly more modulated by simultaneity during SI than during SH, and with a strong trend for such a dissociation in the left putamen. In both the left and right putamen, there was no significant global difference in response to the 2 tasks. The left putamen has been implicated in translation (Klein et al. 1995; Price et al. 1999), in second-language single-word repetition (Klein et al. 1994), and more recently it has been shown to be engaged when multilingual individuals employ a language that is not mastered in a native-like fashion (Abutalebi et al. 2012a). Although our participants produced speech in a language in which they had an extremely high proficiency, the results can be reconciled by considering the putamen as the source of ongoing suppression of speech production in the inappropriate language during SI, that is, while simultaneously hearing sentences in that language. The pragmatics of speech are such that when hearing one language it is normal to respond with speech in that same language and, in naive individuals, suppressing this prepotent response may call upon the putamen substantially more than in experienced interpreters. Thus, we suggest that the role of the putamen in polyglot language control can be thought to be the suppression of

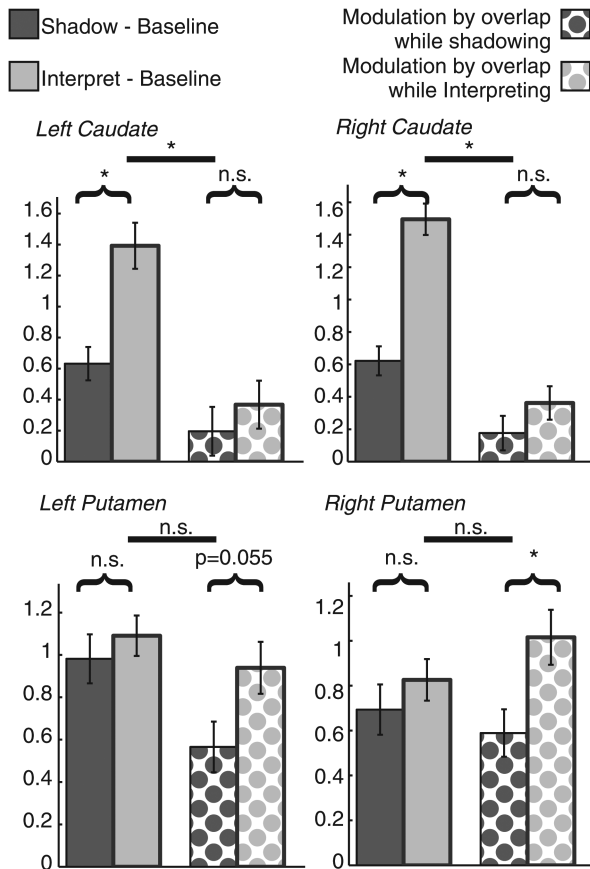


Figure 4. Bar charts showing contrast estimates of left and right caudate and putamen, for contrasts with baseline and for modulation by overlap for shadowing and interpretation. Values are averaged over participants and represent the first eigenvalue of the T-statistic in the regions defined as showing a significant effect for the contrast of interpretation–shadowing (the caudate nuclei) and modulation by overlap (the putamen, bilaterally). Error bars indicate ± 1 standard error of the mean, corrected to be appropriate for within-subjects comparisons (Loftus and Masson 1994).

a contextually dominant language. Taken together, our results suggest a dissociation between the roles of the caudate and the putamen in multilingual control.

The Central Role of the Dorsal Striatum in SI

A striking aspect of these results is the recruitment of the dorsal striatum (the caudate nucleus and the putamen). Although the basal ganglia are often discussed in terms of their role in motor behavior, they play a central role in circuits known to subservise multiple, nonmotoric aspects of cognition such as attention, learning and memory, and executive functions (Saint-Cyr 2003). It has been argued that the basal ganglia and cerebellum interact in loops with cortical regions to select and refine cortical patterns, essentially subtending the ability to control action (Houk et al. 2007).

The linguistic roles of the basal ganglia have received some concerted attention in the past (Lieberman 2000), and there is increasing interest in basal ganglia language functions (Kotz et al. 2009) particularly with regard to multilingualism (e.g., Friederici 2006; Abutalebi and Green 2007, 2008; Abutalebi et al. 2012a; Wang et al. 2012; Zou et al. 2012). The basal ganglia have previously been found to be involved in tasks requiring bilingual language control, such as single-word

translation (Klein et al. 1995; Price et al. 1999), second-language word repetition (Klein et al. 1994, 2006), consecutive interpretation (Lehtonen et al. 2005), and language switching (Price et al. 1999; Abutalebi and Green 2008; Garbin et al. 2010). Results have been equivocal as to the precise roles of these structures. Some authors have reported that language switching recruits the caudate nucleus (Abutalebi and Green 2008; Garbin et al. 2010), whereas others have reported involvement of both caudate and putamen (Price et al. 1999). Similarly, studies on translation have given rise to variable results, with one showing recruitment of putamen (Klein et al. 1995) and, another, external globus pallidus (Lehtonen et al. 2005). Second-language word repetition has been reported to recruit the left caudate (Klein et al. 2006) and the left putamen (Klein et al. 1994). Other authors have not reported basal ganglia involvement during language switching tasks (Wang et al. 2007, 2009) or, as previously noted, during SI (Rinne et al. 2000; Tommola et al. 2000). Nevertheless, a central role for the striatum in bilingual language control is supported by case studies of subcortical polyglot aphasia, which have reported that damage to the putamen or caudate in multilingual individuals can result in involuntary language mixing or switching (Hervais-Adelman et al. 2011; Abutalebi et al. 2012a).

The caudate nucleus and putamen are engaged by SI in our study, and, frequently, by tasks explicitly demanding multilingual control in previous work (see above). In nonlinguistic domains, the caudate is typically described as being involved in the selection of appropriate behavior as a function of predicted outcomes, while the putamen is associated with the implementation and co-ordination of such behavior (Grahn et al. 2008). A compatible distinction between the caudate and putamen is proposed by Ali et al. (2010), who examined the neural correlates of a Stroop and Simon task in fMRI. They concluded that the left caudate was implicated in the inhibition of action plans triggered by incongruent words (in the Stroop task), while the left putamen was implicated in response change in both the Stroop (word interference) and Simon tasks (spatial interference). These distinctions between the caudate and putamen are consistent with our results and, together with them, help to reconcile the existing literature on basal ganglia structures in multilingualism.

We propose that the caudate nucleus is engaged in relatively high-level monitoring and controlling of “language set” selection, controlling which lexico-semantic system is “live” at a given moment. This concurs with the notion that the caudate is involved in determining context-appropriate behavior (e.g., Graybiel 1995; Chee 2006). In contrast, we propose that the putamen is engaged in moment-to-moment language output control at what can be considered a lower level for example in inhibiting the nontarget language in order to favor access to articulatory motor representations required for articulation in the appropriate language. The observed complementary contributions of the caudate and of the putamen during language control reflect an organizing principle of subcortical involvement in the coordination of cognition, action, and learning—the caudate being implicated in planning and the putamen in execution of actions.

The bilateral nature of the findings we report here expand those of several previous reports cited above, in which the left striatum is principally implicated in tasks requiring multilingual language control (Klein et al. 1994, 1995, 2006; Price et al.

1999; Lehtonen et al. 2005; Abutalebi and Green 2008; Garbin et al. 2010). There is also some indication from direct electrical stimulation that the dominant striatum (typically the left striatum in right-handed individuals) is involved in the control of speech production, with the putamen more implicated in the coordination of speech articulation, and the caudate involved in inhibition and selection, in monolingual patients (Gil Robles et al. 2005). A recent meta-analysis of neuroimaging studies of language switching, however, indicates that both right and left striatal structures are implicated in language switching (Luk et al. 2012). Consistent with this, our data indicate that the bilateral caudate nuclei and putamen are implicated in SI. Our bilateral findings may be due to the fact that owing to the large number of participants that we tested, we had relatively higher statistical power, or it may be because in contrast to most previous reports, we used a task that makes large, ongoing (as opposed to momentary) demands on language control. The findings herein reported demonstrate bilateral striatal involvement in a demanding, continuous language control task. (It is of particular interest to note that when including 7 left-handed participants in our analysis [see Supplementary Tables 1–4 and Supplementary Figures 1–3], the lateralization of these results does not qualitatively change, suggesting that the language control functions of these structures may not be substantially affected by handedness. This finding merits further investigation in the future.)

Our results provide new insights into the profound overlap between the neural bases of extreme language control and those of domain-general control of cognition and action. Indeed, recent evidence suggests that experienced simultaneous interpreters display enhanced cognitive flexibility compared even with bilingual individuals (Yudes et al. 2011; Stavrakaki et al. 2012). The recruitment of similar fronto-subcortical-cerebellar circuits during language and executive control provides powerful evidence that the continuous demands of language control in the multilingual brain, and associated experience-dependent plasticity, could underlie the nonlinguistic, executive advantages that have been observed in bilingual individuals, advantages that may also be protective in defying challenges posed by aging and even disease.

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>.

Funding

This work was supported by the Swiss National Science Foundation (grant nos. PP00P3_133701 and 320030_122085).

Notes

We thank Alain Dagher for helpful comments on the manuscript, as well as 3 anonymous reviewers for their thoughtful suggestions. We also thank Frédéric Grouiller and Maria Pefkou for their assistance with data collection, and Sophie Hengl, Violeta Seretan, Magdalena Olivera-Tovar, and Carmen Delgado Luchner for their efforts in rating the participants' output. *Conflict of Interest:* None declared.

References

- Abutalebi J, Annoni JM, Zimine I, Pegna AJ, Seghier ML, Lee-Jahnke H, Lazeyras F, Cappa SF, Khateb A. 2008. Language control and lexical competition in bilinguals: an event-related fMRI study. *Cereb Cortex*. 18:1496–1505.
- Abutalebi J, Della Rosa PA, Castro Gonzaga AK, Keim R, Costa A, Perani D. 2012a. The role of the left putamen in multilingual language production. *Brain Lang*. 125:307–315.
- Abutalebi J, Della Rosa PA, Green DW, Hernandez M, Scifo P, Keim R, Cappa SF, Costa A. 2012b. Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb Cortex*. 22:2076–2086.
- Abutalebi J, Green D. 2007. Bilingual language production: the neurocognition of language representation and control. *J Neurolinguistics*. 20:242–275.
- Abutalebi J, Green D. 2008. Control mechanisms in bilingual language production: neural evidence from language switching studies. *Lang Cogn Process*. 23:557–582.
- Abutalebi J, Tettamanti M, Perani D. 2009. The bilingual brain: linguistic and non-linguistic skills. *Brain Lang*. 109:51–54.
- Ali N, Green DW, Kherif F, Devlin JT, Price CJ. 2010. The role of the left head of caudate in suppressing irrelevant words. *J Cogn Neurosci*. 22:2369–2386.
- Aron AR. 2008. Progress in executive-function research: from tasks to functions to regions to networks. *Curr Dir Psychol Sci*. 17:124–129.
- Badre D, Wagner AD. 2007. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*. 45:2883–2901.
- Bialystok E, Craik FI, Luk G. 2012. Bilingualism: consequences for mind and brain. *Trends Cogn Sci*. 16:240–250.
- Boersma P, Weenink D. 2011. Praat: doing phonetics by computer. Version 5.2.23.
- Bookheimer S. 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci*. 25:151–188.
- Botvinick MM, Cohen JD, Carter CS. 2004. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci*. 8:539–546.
- Brett M, Anton JL, Valabregue R, Poline J-B. 2002. Region of interest analysis using an SPM toolbox [abstract]. In: 8th International Conference on Functional Mapping of the Human Brain. Sendai, Japan. Available on CD-ROM in NeuroImage, Vol 16, No 2.
- Chee MW. 2006. Dissociating language and word meaning in the bilingual brain. *Trends Cogn Sci*. 10:527–529.
- Costa A. 2005. Lexical access in bilingual production. In: Kroll J, De Groot AMB, editors. *Handbook of bilingualism: psycholinguistic approaches*. Oxford, UK: Oxford University Press. p. 308–325.
- Costa A, Hernandez M, Costa-Faidella J, Sebastian-Galles N. 2009. On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition*. 113:135–149.
- Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, Aso T, Urayama S, Fukuyama H, Stockton K et al. 2006. Language control in the bilingual brain. *Science*. 312:1537–1540.
- Dapretto M, Bookheimer SY. 1999. Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron*. 24:427–432.
- Diamond J. 2010. The benefits of multilingualism. *Science*. 330:332–333.
- Dijkstra T, van Heuven WJB. 2002. The architecture of the bilingual word recognition system: from identification to decision. *Biling Lang Cogn*. 5:175–197.
- Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE. 2008. A dual-networks architecture of top-down control. *Trends Cogn Sci*. 12:99–105.
- Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT. 2009. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp*. 30(9):2907–2926.
- Ericsson KA, Charness N, Feltovich PJ, Hoffman RR. 2006. *The Cambridge handbook of expertise and expert performance*. New York (NY): Cambridge University Press.
- Fabbro F, editor. 1999. *The neurolinguistics of bilingualism: an introduction*. Hove: Psychology Press.

- Friederici AD. 2006. What's in control of language? *Nat Neurosci.* 9:991–992.
- Garbin G, Costa A, Sanjuan A, Forn C, Rodriguez-Pujadas A, Ventura N, Belloch V, Hernandez M, Avila C. 2011. Neural bases of language switching in high and early proficient bilinguals. *Brain Lang.* 119:129–135.
- 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.
- Gilbert SJ, Spengler S, Simons JS, Steele JD, Lawrie SM, Frith CD, Burgess PW. 2006. Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J Cogn Neurosci.* 18:932–948.
- Gil Robles S, Gagnol P, Capelle L, Mitchell MC, Duffau H. 2005. The role of dominant striatum in language: a study using intraoperative electrical stimulations. *J Neurol Neurosurg Psychiatry.* 76:940–946.
- Golestani N, Price CJ, Scott SK. 2011. Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *J Neurosci.* 31:4213–4220.
- Grahn JA, Parkinson JA, Owen AM. 2008. The cognitive functions of the caudate nucleus. *Prog Neurobiol.* 86:141–155.
- Graybiel AM. 1995. Building action repertoires: memory and learning functions of the basal ganglia. *Curr Opin Neurobiol.* 5:733–741.
- Green D. 1998. Mental control of the bilingual lexico-semantic system. *Biling Lang Cogn.* 1:67–81.
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. 1999. “Sparse” temporal sampling in auditory fMRI. *Hum Brain Mapp.* 7:213–223.
- Hervais-Adelman A, Moser-Mercer B, Golestani N. 2011. Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. *Front Psychol.* 2:234.
- Hilchey MD, Klein RM. 2011. Are there bilingual advantages on non-linguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon Bull Rev.* 18:625–658.
- Hoshino N, Thierry G. 2011. Language selection in bilingual word production: electrophysiological evidence for cross-language competition. *Brain Res.* 1371:100–109.
- Houk JC, Bastianen C, Fansler D, Fishbach A, Fraser D, Reber PJ, Roy SA, Simo LS. 2007. Action selection and refinement in subcortical loops through basal ganglia and cerebellum. *Philos Trans R Soc Lond B Biol Sci.* 362:1573–1583.
- Huster RJ, Westerhausen R, Pantev C, Konrad C. 2010. The role of the cingulate cortex as neural generator of the N200 and P300 in a tactile response inhibition task. *Hum Brain Mapp.* 31:1260–1271.
- Kello CT, Kawamoto AH. 1998. Runword: an IBM-PC software package for the collection and acoustic analysis of speeded naming responses. *Behav Res Methods Instrum Comput.* 30:371–383.
- Khateb A, Abutalebi J, Michel CM, Pegna AJ, Lee-Jahnke H, Annoni JM. 2007. Language selection in bilinguals: a spatio-temporal analysis of electric brain activity. *Int J Psychophysiol.* 65:201–213.
- Klein D, Milner B, Zatorre RJ, Meyer E, Evans AC. 1995. The neural substrates underlying word generation—a bilingual functional-imaging study. *Proc Natl Acad Sci USA.* 92:2899–2903.
- Klein D, Watkins KE, Zatorre RJ, Milner B. 2006. Word and nonword repetition in bilingual subjects: a PET study. *Hum Brain Mapp.* 27:153–161.
- Klein D, Zatorre RJ, Milner B, Meyer E, Evans AC. 1994. Left putaminal activation when speaking a second language: evidence from PET. *Neuroreport.* 5:2295–2297.
- Kotz SA, Schwartze M, Schmidt-Kassow M. 2009. Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. *Cortex.* 45:982–990.
- Kroll JF, Bobb SC, Misra M, Guo T. 2008. Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol (Amst).* 128:416–430.
- Lehtonen MH, Laine M, Niemi J, Thomsen T, Vorobyev VA, Hugdahl K. 2005. Brain correlates of sentence translation in Finnish-Norwegian bilinguals. *Neuroreport.* 16:607–610.
- Lieberman P. 2000. *Human language and our reptilian brain: the subcortical bases of speech, syntax, and thought.* Cambridge, MA: Harvard University Press.
- Loftus GR, Masson MEJ. 1994. Using confidence intervals in within-subjects designs. *Psychonomic Bull Rev.* 1:476–490.
- Luk G, Green DW, Abutalebi J, Grady C. 2012. Cognitive control for language switching in bilinguals: a quantitative meta-analysis of functional neuroimaging studies. *Lang Cogn Process.* 27:1479–1488.
- Moreno EM, Rodriguez-Fornells A, Laine M. 2008. Event-related potentials (ERPs) in the study of bilingual language processing. *J Neuro-linguistics.* 21:477–508.
- Moser B. 1978. Simultaneous interpretation: a hypothetical model and its practical application. In: Gerver D, Sinaiko HW, editors. *Language, interpretation and communication.* New York: Plenum Press. p. 353–368.
- Moser-Mercer B. 2000. Simultaneous interpreting: cognitive potential and limitations. *Interpreting.* 5:83–94.
- Moser-Mercer B, Frauenfelder UH, Casado B, Künzli A. 2000. Searching to define expertise in interpreting. In: Dimitrova BE, Hyltenstam K, editors. *Language processing and simultaneous interpreting: interdisciplinary perspectives.* Amsterdam: John Benjamins. p. 107–132.
- Nachev P, Kennard C, Husain M. 2008. Functional role of the supplementary and pre-supplementary motor areas. *Nat Rev Neurosci.* 9:856–869.
- Nieuwenhuis S, Yeung N, Van Den Wildenberg W, Ridderinkhof KR. 2003. Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cogn Affect Behav Neurosci.* 3:17–26.
- Peeva MG, Guenther FH, Tourville JA, Nieto-Castanon A, Anton JL, Nazarian B, Alario FX. 2010. Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage.* 50:626–638.
- Peschke C, Ziegler W, Kappes J, Baumgaertner A. 2009. Auditory-motor integration during fast repetition: the neuronal correlates of shadowing. *Neuroimage.* 47:392–402.
- Petersen SE, Posner MI. 2012. The attention system of the human brain: 20 years after. *Annu Rev Neurosci.* 35:73–89.
- Picard N, Strick PL. 1996. Motor areas of the medial wall: a review of their location and functional activation. *Cereb Cortex.* 6:342–353.
- Price CJ, Green DW, von Studnitz R. 1999. A functional imaging study of translation and language switching. *Brain.* 122(Pt 12):2221–2235.
- Quaresima V, Ferrari M, van der Sluijs MCP, Menssen J, Colier W. 2002. Lateral frontal cortex oxygenation changes during translation and language switching revealed by non-invasive near-infrared multi-point measurements. *Brain Res Bull.* 59:235–243.
- Riecker A, Mathiak K, Wildgruber D, Erb M, Hertrich I, Grodd W, Ackermann H. 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology.* 64:700–706.
- Rinne JO, Tommola J, Laine M, Krause BJ, Schmidt D, Kaasinen V, Teras M, Sipila H, Sunnari M. 2000. The translating brain: cerebral activation patterns during simultaneous interpreting. *Neurosci Lett.* 294:85–88.
- Rodriguez-Fornells A, Balaguer RDD, Münte TF. 2006. Executive control in bilingual language processing. *Lang Learn.* 56:133–190.
- Rushworth M, Walton M, Kennerley S, Bannerman D. 2004. Action sets and decisions in the medial frontal cortex. *Trends Cogn Sci.* 8:410–417.
- Saint-Cyr JA. 2003. Frontal-striatal circuit functions: context, sequence, and consequence. *J Int Neuropsychol Soc.* 9:103–127.
- Schweizer TA, Ware J, Fischer CE, Craik FI, Bialystok E. 2012. Bilingualism as a contributor to cognitive reserve: evidence from brain atrophy in Alzheimer's disease. *Cortex.* 48(8):991–996.
- Stavrakaki S, Megari K, Kosmidis MH, Apostolidou M, Takou E. 2012. Working memory and verbal fluency in simultaneous interpreters. *J Clin Exp Neuropsychol.* 34(6):624–633.
- Tommola J, Laine M, Sunnari M, Rinne JO. 2000. Images of shadowing and interpreting. *Interpreting.* 5:147–169.
- van Heuven W, Dijkstra T. 2010. Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Res Rev.* 64:104–122.
- Wang X, Wang YY, Jiang T, Wang YZ, Wu CX. 2012. Direct evidence of the left caudate's role in bilingual control: an intra-operative electrical stimulation study. *Neurocase.* 19(5):462–469.

- Wang Y, Kuhl PK, Chen C, Dong Q. 2009. Sustained and transient language control in the bilingual brain. *Neuroimage*. 47:414–422.
- 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.
- Yudes C, Macizo P, Bajo T. 2011. The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Front Psychol*. 2:309.
- Zou L, Ding G, Abutalebi J, Shu H, Peng D. 2012. Structural plasticity of the left caudate in bimodal bilinguals. *Cortex*. 48:1197–1206.