

Where, When and Why Brain Activation Differs for Bilinguals and Monolinguals during Picture Naming and Reading Aloud

Öiwi Parker Jones¹, David W. Green², Alice Grogan³, Christos Pliatsikas⁴, Konstantinos Filippopolitis¹, Nilufa Ali⁵, Hwee Ling Lee⁶, Sue Ramsden¹, Karine Gazarian¹, Susan Prejawa¹, Mohamed L. Seghier¹ and Cathy J. Price¹

¹Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, UK, ²Cognitive, Perceptual and Brain Sciences, University College London, London WC1E 6BT, UK, ³Department of Human Communication Science, University of Sheffield, Sheffield S10 2TA, UK, ⁴Centre for Integrative Neuroscience and Neurodynamics, University of Reading, Reading RG6 6AH, UK, ⁵Department of Psychology, Anglia Ruskin University, Cambridge CB1 1PT, UK and ⁶Cognitive Neuroimaging Group, Max-Planck Institute for Biological Cybernetics, Tübingen 72076, Germany

Address correspondence to Öiwi Parker Jones, Wellcome Trust Centre for Neuroimaging, 12 Queen Square, London WC1N 3BG, UK.
Email: o.parkerjones@fil.ion.ucl.ac.uk

Using functional magnetic resonance imaging, we found that when bilinguals named pictures or read words aloud, in their native or nonnative language, activation was higher relative to monolinguals in 5 left hemisphere regions: dorsal precentral gyrus, pars triangularis, pars opercularis, superior temporal gyrus, and planum temporale. We further demonstrate that these areas are sensitive to increasing demands on speech production in monolinguals. This suggests that the advantage of being bilingual comes at the expense of increased work in brain areas that support monolingual word processing. By comparing the effect of bilingualism across a range of tasks, we argue that activation is higher in bilinguals compared with monolinguals because word retrieval is more demanding; articulation of each word is less rehearsed; and speech output needs careful monitoring to avoid errors when competition for word selection occurs between, as well as within, language.

Keywords: bilingualism, control, fMRI, frequency, picture naming, reading

Introduction

Language is a core cognitive function and so it is essential to understand how the brain adapts to using 2 or more languages and what this adaptation means. For example, do the advantages of being bilingual come from uniquely bilingual processing or result from increased demands on resources that also support monolingual processing? To address this question, it is important to compare bilingual and monolingual brain activation when bilinguals are tested in only one of their languages because switching between languages can induce additional processing requirements (Grosjean 1998, 2001). In the present study, we used functional Magnetic Resonance Imaging (fMRI) to investigate whether neuronal activation differs in bilinguals and monolinguals during picture naming and reading aloud when only one language is in use.

To our knowledge, no previous fMRI study has compared speech production in bilinguals relative with monolinguals when bilinguals are tested in only one of their languages within an experimental session. We therefore start by considering the relevance of previous studies that have reported differences in bilingual and monolingual brain activation when bilinguals were required to use 2 languages within an experimental session. In the fMRI and functional Near-Infrared Spectroscopy (fNIRS) experiments reported by Kovelman and colleagues, activation in the left inferior frontal cortex was greater in Spanish-English bilinguals than in English monolinguals when

participants made syntactic judgments (Kovelman, Baker, et al. 2008) or semantic judgments (Kovelman, Shalinsky, et al. 2008). The authors interpreted their results in terms of a unique “bilingual signature” which represents an advantageous “language processing potential not recruited in monolingual brains” (Kovelman, Baker, et al. 2008; also see Kovelman, Shalinsky, et al. 2008). In a further study, Kovelman et al. (2009) reported greater activation in the left temporoparietal cortex when bilingual English-American sign language users with normal hearing responded to pictures in a dual language context (either by simultaneously naming and signing each one or by naming one set of pictures and signing the next) compared with a single language context (in which only one language was used over a block of trials).

Increased activation in left inferior frontal and superior temporal areas has also been reported for Spanish-Catalan bilinguals compared with Spanish monolinguals performing a speeded response task for words in Spanish that begin with a consonant or a vowel (i.e., responding with different buttons for each), while at the same time suppressing a response for pseudowords or Catalan words (Rodríguez-Fornells et al. 2002). The Catalan words interfered when the participants were Spanish-Catalan bilinguals but not when the participants were Spanish monolinguals. The authors therefore suggested that increased activation was related to the control of “lexical” interference. In a subsequent study (Rodríguez-Fornells et al. 2005), these authors found increased activation in a left middle prefrontal region for bilinguals relative to monolinguals when “phonological” interference was introduced in a task that required participants to decide if the first letter of an object’s name was a German vowel. On half the trials, there was a mismatch between the correct response in German and Spanish (e.g., *Erdbeere-fresa* for strawberry) which resulted in response interference when the participants were German-Spanish bilinguals but not when the participants were German monolinguals.

Although the above studies have highlighted increased activation for bilinguals relative to monolinguals in similar left inferior frontal and temporoparietal areas, it is not possible to compare the exact location of the effects in the different tasks tested because no anatomical details are provided in the fNIRS studies by Kovelman, Shalinsky, et al. (2008) and Kovelman et al. (2009). Moreover, because speech production activation in bilinguals and monolinguals has only been compared in the dual language context for bilinguals, we do not know how the effects would generalize to a single language context.

The first aim of the current study was to establish “whether” brain activation differed for bilinguals and monolinguals during picture naming and reading aloud when the bilinguals are tested in a single language context, either in their native language or in a foreign language (but not both on the same day). There are 2 reasons why we expected activation to be higher in bilinguals compared with monolinguals during picture naming and reading aloud in a native or nonnative language. First, the interference control hypothesis proposes that by knowing the name of a concept in 2 or more languages, bilinguals must selectively activate the target language while minimizing competition for word selection from translation equivalents in nontarget languages (Green 1986, 1998). Learning and using another language increases interference and places additional demands on the mechanisms that control interference (Green 1986, 1998; Grosjean 1992, 2001; Rodriguez-Fornells et al. 2002, 2005; Abutalebi and Green 2007). Second, the reduced frequency hypothesis proposes that words in each of a bilingual’s languages are effectively used less than the same words in a monolingual speaker’s language because monolinguals always use the same language to express themselves whereas bilinguals split their time between 2 or more languages. Consequently, the frequency of a word in each of the bilingual’s languages effectively lags behind the frequency of the same word in a monolingual’s language. As lower frequency words are more difficult to produce, bilingual word production will be less efficient than monolingual word production (for behavioral studies, see Mägiste 1979; Ransdell and Fischler 1987; Gollan et al. 2002, 2005, 2008; Gollan and Acenas 2004; Ivanova and Costa 2008; Pyers et al. 2009), despite psycholinguistic and neuroimaging research showing that the languages of a bilingual are both active even when only one language is in use (van Heuven and Dijkstra 2010; Wu and Thierry 2010).

Crucially, our 2 hypotheses—interference control and reduced frequency—are not mutually exclusive. To the contrary, the control of interference from translation equivalents may utilize the same resources as the control of interference from synonyms within a single language (e.g., sofa and couch) which also makes word retrieval more difficult (e.g., Jescheniak and Schriefers 1998; Peterson and Savoy 1998). Furthermore, producing low- versus high-frequency words may also require processes that resolve interference from competing possibilities. If this is indeed the case, then the brain regions identified for bilingual versus monolingual picture naming may be the same as the brain regions responsive to low- versus high-frequency picture naming in monolinguals. Alternatively, if the control of competition between languages requires different processing resources than the control of competition within a single language, then the brain regions identified with such control would be distinct from those associated with low- versus high-frequency word processing within a language.

The second aim of our study was therefore to identify “where” activation is higher in bilinguals (speaking in a single language context) than monolinguals and to compare the location of our effects with the location of activation reported in previous literature by 1) Graves et al. (2007) for low- versus high-frequency picture naming in the native language; 2) Papoutsis et al. (2009) for articulating nonwords with low- versus high-frequency syllables; and 3) Rodriguez-Fornells et al. (2002, 2005) for bilinguals versus monolinguals when bilinguals were tested in a dual language context that required them to suppress lexical or phonological interference from words in

the nontarget language. Details of these a priori effects are listed in Table 2.

Having determined “whether” and “where” the effect of bilingualism is observed, our third aim was to consider “when,” in the multistage speech production processing stream, the effect of bilingualism arose. Our hypothesis was that bilingualism and low-frequency word processing might increase the demands on 1) word retrieval because the links between semantics and phonology will be weaker in less familiar words; 2) articulation because each word is less rehearsed; and 3) control mechanisms that suppress competition from nontarget words. To dissociate the processing stage when bilingual and monolingual activation diverged, we engaged all our participants in a range of tasks that differentially tapped processing related to word recognition, word retrieval, and articulation (see Fig. 1).

Finally, we tested whether activation in the identified areas correlated with the participants’ ability to control conflicting verbal information as assessed outside the scanner with the Stroop task. Participants named the color (i.e., the hue) of a stimulus that was either a string of XXXs (=neutral or nonconflict trial) or the written name of a color (=conflict trials because the meaning/phonology of the word competes with that of the hue); see Long and Prat (2002). We used the conflict ratio for response time CR_{RT} (Green et al. 2010) as our index of control. CR_{RT} is the response time difference between conflict trials (CT) and neutral trials (NT), divided by NT:

$$CR_{RT} = \frac{CT - NT}{NT}$$

The conflict ratio for response time indicates how well interference has been controlled for correct responses. If the ratio was low, then interference was low and there was a higher level of control. Alternatively, if the ratio was high, then interference was high and there was a lower level of

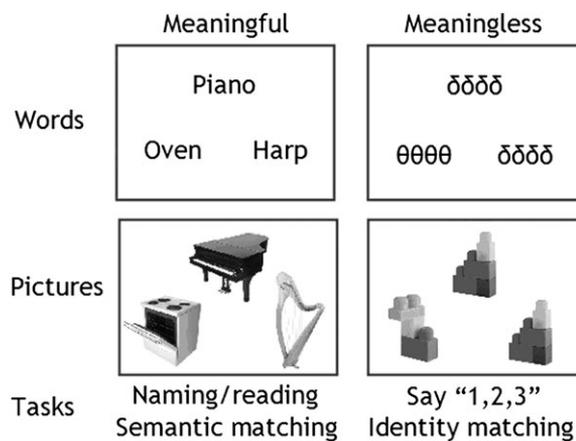


Figure 1. Paradigm. Our experimental paradigm was designed to tease apart activation related to word retrieval, word recognition, and articulation. In all trials, 3 stimuli were simultaneously presented as a “triad,” with 1 stimulus above and 2 stimuli below. Word retrieval was assessed during picture naming (condition 1) and reading aloud (condition 2). Articulation without word retrieval involved articulating “1,2,3” in response to pictures of unfamiliar nonobjects (condition 3) or unfamiliar (and meaningless) strings of Greek letters (condition 4). Word recognition and semantic processing were assessed during semantic decisions (e.g., matching Piano to Harp rather than Oven) on pictures (condition 5) or words (condition 6), using a finger press response. Perceptual processing was assessed during a physical identity match on the meaningless Greek letter strings (condition 7) or pictures of nonobjects (condition 8), using a finger press response.

control. We therefore associated a positive correlation between the conflict ratio and activation in the identified areas with interference. In contrast, we associated a negative correlation between the conflict ratio and activation in the identified areas with a mechanism that controls interference.

In summary, we investigated whether brain activation differs for bilingual versus monolingual speech production when bilinguals were tested in a single language context that required use of just one of their languages (either native or nonnative). We then examined “where” the effects were located relative to previous studies of the control of interference in bilinguals and studies of word frequency effects in monolinguals. This allowed us to test whether the effect of learning 2 languages results in activation that is unique to bilinguals or whether it is also observed in monolinguals processing low- versus high-frequency words. By including a range of tasks that tapped word recognition, word retrieval, articulation, and the control of interference, we were able to ascertain “when” in the processing stream differences in bilingual and monolingual activation arise and so consider, with reference to other studies, “why” such differences arise.

Materials and Methods

The study was approved by the National Hospital for Neurology and Institute of Neurology Joint Ethics Committee.

Participant Selection

A total of 67 participants were included. All were neurologically typical, right handed, fluent English speakers with normal or corrected to normal vision. The bilinguals ($n = 31$) spoke English as a nonnative language but were all resident in the UK with high English proficiency. Their native languages were European ones with Latin-based scripts ($n = 10$) or Greek ($n = 21$); for details, see Table 1. We further divided the Greek-English bilinguals into those who were scanned in English only ($n = 10$) and those who were scanned in both English and Greek ($n = 11$). This resulted in 3 bilingual groups that we refer to as Groups 2a, 2b, and 2c (see Table 1). To ensure that bilinguals were tested in a single language context, all word stimuli and instructions were only given in the language being tested, and Greek participants were tested in English and Greek on separate days.

The monolinguals ($n = 36$, referred to as Group 1) did not use a second language at home or at work. Although we were not able to quantify the degree to which they had been exposed to other languages, we note that such experience would not be able to explain the highly significant group differences that we report in this study.

Participant Screening

To ensure that all our bilinguals had high English proficiency, they completed a lexical decision test from the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay et al. 1992). We also recorded responses for letter and category fluency (Grogan et al. 2009) and used a color-word version of the Stroop task (Stroop 1935) to assess the control of verbal interference in both the native language and English. The results (see Table 1) confirmed that all bilinguals had an extensive vocabulary in English and rapid written word recognition. For example, all but one Greek subject (from Group 2b in Table 1) were able to recognize English words and reject pseudowords in the PALPA decision task with a high level of accuracy (>80%) and speed (<1 s). When the outlier participant was excluded, there were no significant differences in the 3 bilingual groups on the speed or accuracy of their lexical decisions (analysis of variance, ANOVA with 3 groups). However, as expected, the bilinguals were slightly slower and less accurate than the monolingual English participants for both correctly accepting words ($t_{41} = -2.93$, $P < 0.05$ for RTs; $t_{40.56} = 4.84$, $P < 0.05$ for accuracy) and for correctly rejecting nonwords [$t_{39.03} = -4.63$, $P < 0.05$ for RTs; $t_{40.42} = 3.86$, $P < 0.05$ for accuracy].

Table 1

Participant groups

	Groups			
	Monolingual	Bilingual		
	1	2a	2b	2c
<i>n</i> (females)	36 (24)	10 (3)	10 (4)	11 (6)
Mean age	39.7	43.0	24.6	28.4
Age range	18–73	27–69	21–69	23–58
L1	Eng	GIDC	Gk	Gk
Number of languages				
Mean (range) years	1 (1–1)	4.3 (2–8)	2.5 (2–4)	3 (2–4)
Age of English acquisition				
Mean (range) years	Native	10.3 (1–15)	8.8 (3–14)	9 (3–12)
In-scanner Eng/Gk accuracy (%)				
Naming	96.3/na	90.2/na	84.8/na	89.0/88.3
Reading aloud	99.7/na	98.7/na	99.3/na	97.0/98.2
Semantics pictures	91.6/na	88.8/na	93.6/na	90.1/92.4
Semantics words	92.4/na	87.3/na	85.9/na	86.3/92.4
In-scanner Eng/Gk RT (s)				
Semantics pictures	1.77/na	1.90/na	1.70/na	1.66/1.86
Semantics words	1.69/na	1.99/na	2.01/na	2.01/1.95
Out of scanner PALPA ^a	$n = 13$	$n = 9$	$n = 10$	$n = 11$
Mean accuracy (%) real words	97.7	93.9	88.5	90.0
Range accuracy words	91–100	85–100	72–98	82–100
Mean accuracy (%) nonwords	97.2	91.4	89.5	92.5
Range accuracy nonwords	89–100	86–100	68–100	88–100
RT (ms) real words	758	855	826	858
RT (ms) nonwords	851	1117	1080	1096
Out of scanner FLUENCY	$n = 16$	$n = 9$	$n = 9$	$n = 10$
Mean category and letter	19.8	17.7	14.4	16.65
Out of scanner STROOP (Eng/L1)	$n = 12$	$n = 10$	$n = 7$	$n = 10$
Mean neutral correct (%)	92/92	89/87	100/100	98/99
Mean conflict correct (%)	92/92	83/74	99/86	97/85
Mean neutral RT (ms)	752/752	801/804	897/888	895/943
Mean conflict RT (ms)	880/880	967/909	1047/1092	974/1057

Note: English (Eng); German, Italian, Dutch or Czech (GIDC); Greek (Gk); First language (L1); Number (*n*); Not Available (na). Reaction time (RT) refers to correct responses.

^aThirteen controls completed some or all of the out of scanner behavior. One Italian from Group 2a did not complete the out of scanner behavior but self-rated with maximum proficiency score (9 on understanding, speaking, reading, and writing).

We also compared self-ratings for English proficiency (on a scale from 1 to 9, where 1 was low and 9 was high proficiency). The monolinguals (Group 1) rated themselves at ceiling (mean = 9, range = [9,9]). Groups 2a, 2b, and 2c had means of 8.8, 8.3, and 7.1, respectively. In a post hoc Games-Howell analysis of a one-way ANOVA (with Group as the factor), there was a significant difference between the 2 extreme Groups, 1 and 2c ($P < 0.05$), but given the excellent performance of all participants on the PALPA tests, we cannot exclude the possibility that differences in self-ratings reflect differences in confidence.

No group effects were found for category and letter fluency (ANOVA with 3 groups of bilinguals) or for the control of verbal interference during the Stroop task (i.e., CR_{RT} in native language or in English for monolinguals or bilinguals).

Functional Magnetic Resonance Imaging

Experimental Paradigm

Our experimental paradigm was designed to tease apart activation related to word retrieval, word recognition, and articulation. There were 8 different conditions: 4 required a speech production response and 4 required a decision and finger press response. For each type of response, there were 4 types of stimuli: pictures of familiar objects, written object names, pictures of nonobjects, and Greek letters/symbols (see Fig. 1). During the speech production conditions, the participants named the familiar objects, read aloud the written names, and said “1,2,3” to the nonobjects and Greek letters/symbols. During the finger press response conditions, the participants made semantic decisions to the objects and words and perceptual decisions to the nonobjects and Greek letters/symbols.

Areas associated with articulation were those that were more activated in the 4 speech production conditions than the 4 decision/

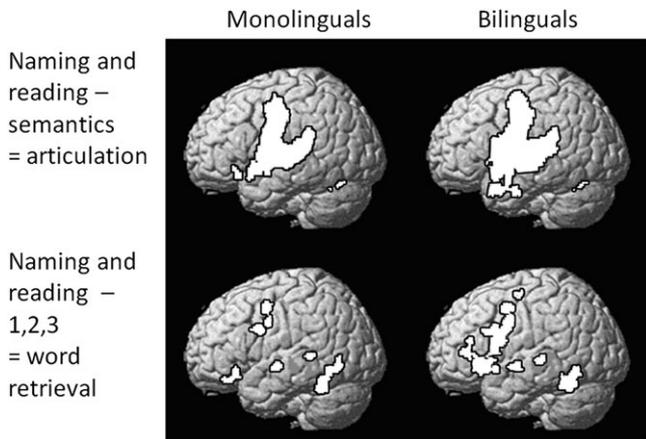


Figure 2. Task effects in monolinguals and bilinguals separately. Greater activation for 1) naming and reading aloud minus semantic conditions (Analysis 1) and 2) naming and reading aloud minus articulating “1,2,3” (Analysis 2). Threshold = $P < 0.001$; $k > 30$ voxels. After a correction for multiple comparison across the whole brain ($P < 0.05$ familywise error corrected), areas that were more activated by naming and reading aloud than by semantic decisions for monolinguals and bilinguals, included bilateral precentral and postcentral gyri, bilateral superior temporal cortices, bilateral cerebellum, the left thalamus, and the supplementary motor cortex. In contrast, areas that were more activated by naming and reading aloud than by articulation (monolinguals and bilinguals) included the left anterior fusiform, anterior cingulate, dorsal premotor cortex, left superior temporal gyrus, left frontal operculum, and bilateral cerebellum.

finger press conditions. Areas associated with word and object recognition were those that were more activated for familiar words and objects than unfamiliar Greek letters and symbols. Areas associated with word retrieval were those more activated for naming and reading aloud than semantic matching on words and objects, after controlling for articulation (i.e., the contrast was [Naming and reading aloud > saying 1,2,3] - [Semantic decisions on objects and words > perceptual decisions on unfamiliar nonobjects and Greek letters/symbols]).

In all trials for all conditions, 3 stimuli were simultaneously presented as a “triad,” with 1 stimulus above and 2 stimuli below (see Fig. 1). Stimuli within a triad were always of the same type (pictures of objects, written object names, meaningless combinations of Greek letters or pictures of meaningless, and unfamiliar nonobjects). The triad configuration was necessary for the semantic and perceptual association matching tasks where a target stimulus above was related to 1 of the 2 stimuli below and participants pressed 1 of 2 response keys with 2 fingers from the same hand to indicate the match. For the articulation conditions, the triad presentation allowed us to keep the visual configuration constant with the semantic and perceptual conditions. However, to avoid interference or priming effects during the articulation conditions, we ensured that stimuli within a triad were semantically and perceptually unrelated to one another. The presentation of 3 naming or reading stimuli at the same time had the added advantage that the presentation rate per item was rapid, thereby maximizing the hemodynamic response per stimulus.

Stimulus Selection

All stimuli were derived from a set of 192 familiar objects with 3–6 letter names in English: 33 had 3 letter names (cat, bus, hat), 65 had 4 letter names (ship, bell, frog, hand), 58 had 5 letter names (teeth, camel, snake), and 36 had 6 letter names (spider, dagger, button). The 192 objects were first divided into 2 different sets of 96 items which we refer to as set A and set B. For example, these were matched by word length (each set had a mean length of 4.5 segments with no statistical difference between sets: $t_{190} = -0.289$, $P = 0.773$). One group of selected participants was presented with set A as written words for reading aloud, set B as pictures for object naming, set B for semantic decisions on words, and set A for semantic decisions on pictures. The other group was presented with set B as written words for reading aloud, set A as pictures for object naming, set A for semantic decisions on words,

and set B for semantic decisions on pictures. Thus, no word or picture was repeated over the 4 runs although 1) each object concept occurred twice (once as a word and once as a picture) and 2) the participants who participated in both English and Greek versions of the paradigm were exposed to the same pictures twice. We tested for the effect of repetition on pictures but we did not expect or find a significant repetition effect because the English and Greek paradigms were conducted on different days with relatively long intervals between testing days (see above). Words and pictures were counterbalanced within and between runs. By using written names and pictures that referred to the same object (e.g., horse), the verbal and nonverbal stimuli were matched for semantic content and associations.

In the naming/reading aloud triads, we minimized the semantic relationship between stimuli such as “lemon” (above), “cow” (lower left), “pipe” (lower right). In the semantic triads, there was a strong semantic relationship between the target and 1 of the lower 2 pictures or words in the triad, for example, “piano” is more semantically related to “harp” than “oven.” We did not include triads where the semantic decision could be made on the basis of perceptual attributes or verbal paradigmatic or syntagmatic associations (e.g., cat and dog, knife and fork, sock and shoe). A pilot study with 8 participants ensured interparticipant agreement on the expected semantic relationship. Further details of the English stimuli can be found in Hu et al. (2010), which reports the same paradigm and provides the full list of object names in the supplementary information.

All word stimuli and instructions were translated into Greek for the Greek paradigm by a proficient Greek-English speaker who did not participate in the study. By using the same stimuli in Greek and English, we ensured that word meaning and picture stimuli were held constant. Any differences in familiarity, word frequency, word length, and other variables might have been problematic for interpreting activation differences in Greek and English (within participant group) but would not affect the interpretation of our findings that activation was higher for bilinguals than monolinguals on English words only (i.e., group differences when the stimuli are held constant).

Procedure

For both the English and the Greek paradigms, each participant participated in 4 scanning runs (sessions) lasting approximately 6 min each, with 2 runs involving the 4 articulation tasks and 2 runs involving the 4 matching, finger press, tasks. Within each run, there were 4 blocks of pictures, 4 blocks of words, 2 blocks of nonobjects, and 2 blocks of unfamiliar Greek letter strings, with 4 triads per block (=12 stimulus items per block). Each block was preceded by 3.6 s of instructions. The instructions in the articulation conditions were “NAME,” “READ,” “123 SYMBOLS,” “123 PICTURES.” The instructions in the match/finger press conditions were “PICTURE-MATCH,” “WORD-MATCH,” “SAME-PICTURE,” “SAME-SYMBOLS.”

Following the 3.6 s of instructions, each triad remained on the screen for 4.32 s followed by 180 ms of fixation, adding up to 18 s for each condition. Blocks of 14.4 s of fixation were interspersed every 2 stimulus blocks. These timing parameters were selected to ensure that the stimulus onset was asynchronized with the slice acquisition which ensured distributed sampling (Veltman et al. 2002). The order of words and pictures was counterbalanced within each run and the order of tasks was counterbalanced across sessions.

To ensure that the task was understood correctly, all participants undertook a short training session with a different set of words and pictures before entering the scanner. For the naming/reading aloud conditions, participants were instructed to name/read aloud the top stimulus first, followed by the bottom left and then the bottom right. For the unfamiliar nonobjects and unfamiliar meaningless letter string conditions, participants were instructed to say 1,2,3 as they systematically viewed the top stimulus, the bottom left stimulus, and the bottom right stimulus. In the semantic matching task, participants were instructed to indicate 1) whether the stimulus on the lower left or lower right was more semantically related to the stimulus above (e.g., is oven or harp most closely related to piano) and 2) for the meaningless triads, whether the lower left or lower right stimulus was visually identical to the one above. Responses were recorded using a button box held either in the left or right hand with the hand of response

constant within each participant and counterbalanced across participants within group. The hand of response was rotated across participants so that the data from the same participants could be used in a future study of bilingual stroke patients who may only be able to use their left or right hand. The specific instructions were to indicate the bottom left stimulus by pressing the index finger on the right hand (or the middle finger on the left hand) and to indicate the bottom right stimulus by pressing the middle finger on the right hand (or the index finger on the left hand).

During training, we emphasized the need to keep the body, head, and mouth as still as possible. In the scanner, stimulus presentation was via a video projector, a front-projection screen, and a system of mirrors fastened to a head coil. Words were presented in lower case Arial and occupied 4.9° (width) and 1.2° (height) of the visual field. Each picture was scaled to take 7.3 × 8.5° of the visual field. Participants' verbal responses, during the articulation conditions, were recorded and filtered using a noise cancellation procedure so that we could monitor accuracy and distinguish correct and incorrect responses in our statistical analysis. However, the recordings were made independently of the presentation script, so the recordings did not contain the stimulus onsets and we were not able to measure naming or reading latencies.

During both the training and the scanning sessions, participants were spoken to in the language they were being tested in. A native Greek speaker who did not participate in the experiment acquired the data for the Greek paradigm and analyzed correct versus incorrect responses.

Data Acquisition

A Siemens 1.5-T Sonata scanner was used to acquire both structural and functional images from all participants (Siemens Medical Systems, Erlangen, Germany). Structural T_1 -weighted images were acquired using a 3D modified driven equilibrium Fourier transform sequence and 176 sagittal partitions with an image matrix of 256 × 224 and a final resolution of 1 mm³ (repetition time (TR)/echo time (TE)/inversion time, 12.24/3.56/530 ms, respectively). Functional T_2^* -weighted echoplanar images with blood oxygen level-dependent contrast comprised 40 axial slices of 2 mm thickness with 1 mm slice interval and 3 × 3 mm in-plane resolution (TR/TE/flip angle = 3600 ms/50 ms/90°, respectively; field of view = 192 mm, matrix = 64 × 64). One hundred and three volumes were acquired per session, leading to a total of 412 volume images across the 4 sessions. As noted above, TR and Stimulus Onset Asynchrony did not match, allowing for distributed sampling of slice acquisition across the experiment (Veltman et al. 2002). To avoid Nyquist ghost artifacts a generalized reconstruction algorithm was used for data processing. After reconstruction, the first 4 volumes of each session were discarded to allow for T_1 equilibration effects.

Preprocessing

Image processing and first-level statistical analyses were conducted using Statistical Parametric Mapping (SPM5; Wellcome Trust Centre for Neuroimaging, London, UK. <http://www.fil.ion.ucl.ac.uk/spm/>) running under Matlab 7 (Mathworks, Sherbon, MA). Each participant's functional volumes were realigned and unwarped (Andersson et al. 2001), adjusting for residual motion-related signal changes. Scans from the different participants were then spatially normalized to Montreal Neurological Institute (MNI) space (voxel size = 2 × 2 × 2 mm³) using unified segmentation/normalization of the structural image after it had been coregistered to the realigned functional images. The normalized functional images were then spatially smoothed with a 6 mm full width half maximum isotropic Gaussian kernel to compensate for residual variability after spatial normalization and to permit application of Gaussian random-field theory for corrected statistical inference. We excluded 3 participants who moved more than 3 mm (1 voxel) in the scanner or when visual inspection of the first-level results indicated movement artifacts (edge effects, activation in ventricles, etc.).

First-Level Analyses

The preprocessed functional volumes of each participant were then submitted to participant-specific fixed-effects analyses, using the general linear model at each voxel. Correct responses for each of the 8 conditions were modeled separately from the instructions and the errors, using event-related delta functions that modeled each trial onset

as an event using condition-specific "stick-functions" having a duration of 4.32 s per trial and a stimulus onset interval of 4.5 s. Each event was convolved with a canonical hemodynamic response function. To exclude low-frequency confounds, the data were high-pass filtered using a set of discrete cosine basis functions with a cutoff period of 128 s. The contrasts of interest at the first level were each of the 8 activation conditions relative to fixation. The appropriate summary or contrast images were then entered into 5 second-level analyses (i.e., a random-effects analysis) to enable inferences at the group level.

Second-Level Analyses

The primary aim of these analyses was to compare bilingual and monolingual activation, while controlling for differences in native versus nonnative language (i.e., English, Greek, or other). We therefore identified activation that differed for all 3 bilingual groups (relative to the monolinguals) irrespective of whether they were responding in their nonnative language (Groups 2a, 2b, 2c) or their native language (Group 2c).

In Analysis 1, we used an ANOVA in SPM with 2 different factors. One factor modeled the 4 different experimental conditions (e.g., naming, reading aloud, semantic decisions on pictures, and semantic decisions on words). The variance for this factor was within subjects. The other factor modeled 5 independent sources of data from the 4 groups performing the English paradigm and the Greek participants performing the Greek paradigm. This second factor was the combination of 2 different nested factors: the effect of monolinguals versus bilinguals performing the English paradigm (Group 1 vs. Groups 2a, 2b, 2c) and the effect of performing the paradigm in L1 (Group 1 and Group 2c in Greek) versus L2 (Group 2a and 2b). In both cases, the contrast is across independent groups; therefore, they can be included in the same analysis. This ensures common sensitivity to all effects because the error variance and degrees of freedom are held constant. In addition to the conditions, we included 2 regressors of no interest. One was the effect of age for each task; the other was in-scanner accuracy on the naming task only (there was very little variance in accuracy on the other tasks).

Analysis 2 was identical to Analysis 1 except that all the semantic conditions were replaced with the articulation conditions (articulating 1,2,3 to unfamiliar nonobjects or unfamiliar letter strings). Analysis 3 correlated brain activation during naming and reading with the ability to control interference in the Stroop task (CR_{RT}, measured outside the scanner). Analysis 4 correlated brain activation with age of acquisition and written word knowledge in L2. Finally, Analysis 5 compared naming and reading activation in those who spoke 1, 2, or 3+ languages. In every condition for all analyses, we only included activation related to correct trials; therefore, all the results reported below represent activations for successful responses.

Statistical Thresholds

For Analysis 1, the statistical threshold was set at $P < 0.05$ familywise error corrected in height for multiple comparisons across 1) the whole brain and 2) regions of interest (spheres of 6 mm radius) centered on coordinates from previous studies of language control in bilinguals and word frequency in monolinguals (see Table 2). For Analyses 2-5, our regions of interest (spheres of 6 mm radius) were centered on the areas identified in Analysis 1.

Results

In-Scanner Behavior

Accuracy for all participants was greater than 80% in all tasks (see Table 1). There were no significant differences between monolinguals and bilinguals (Groups 2a, 2b, 2c in the English paradigm) on either accuracy or response times except during word matching (accuracy: $F_{3,63} = 3.015$, $P < 0.05$; RTs: $F_{3,63} = 7.019$, $P < 0.001$). Tukey post hoc tests showed that monolinguals were faster on word matching than each of the bilingual groups ($P = 0.020$ for group 1 vs. 2a; $P = 0.011$ for group 1 vs. 2b; $P = 0.008$ for group 1 vs. 2c). Likewise, there

Table 2

Predictions

	x	y	z	Factor
Pars triangularis (PTr)				
Kovelman, Baker, et al. (2008)	-48	+38	+4	"Bilingual signature"
Rodriguez-Fornells et al. (2002)	-44	+28	+8	Lexical interference
Graves et al. (2007) ^a	-43	+30	+2	Frequency
Anterior insula (Ins)				
Graves et al. (2007) ^a	-31	+25	+5	Frequency
Pars opercularis (POp)				
Rodriguez-Fornells et al. (2002)	-60	+8	+8	Lexical interference
Papoutsis et al. (2009)	-54	+12	+12	Sublexical frequency
Middle frontal gyrus				
Rodriguez-Fornells et al. (2005)	-40	+36	+32	Phonological interference
Planum temporale (PT)				
Graves et al. (2007) ^a	-52	-39	+20	Frequency
Dorsal precentral gyrus (PrC)				
Graves et al. (2007) ^a	-46	-15	+36	Frequency and word length

^aThese coordinates have been converted from Talairach to MNI space.

were no significant differences for first versus second language (Group 2c in Greek vs. English) except for word matching (accuracy: $t_{10} = 2.566$, $P < 0.05$; RTs: $t_{10} = -2.667$, $P < 0.05$). For details, see Table 1.

The lack of any significant behavioral difference between monolingual and bilingual naming and reading accuracy is consistent with our selection of bilinguals who were highly proficient in English. In essence, we are claiming that the bilinguals were able to perform our relatively easy speech production tasks but they required more brain effort/activation to produce the same outcome (Callan et al. 2004).

Functional Magnetic Resonance Imaging

Prior to reporting group differences in activation, we note that the pattern of activation for word retrieval and articulation was remarkably similar in bilinguals and monolinguals with the exception of the pars opercularis and pars triangularis (POp and PTr) where activation was observed in bilinguals but not monolinguals during naming and reading aloud compared with semantic decisions (Analysis 1) or articulating 1,2,3 (Analysis 2); see Figure 2.

Analysis 1: Activation for Bilinguals and Monolinguals during Naming, Reading Aloud, and Semantic Decisions

Activation that differed for all bilingual groups relative to monolinguals was computed for each task. During naming and reading, bilinguals showed more activation than monolinguals in 5 left hemisphere regions ($P < 0.05$ corrected for multiple comparisons across the whole brain). The first of these was located around the left central sulcus and centered on the posterior surface of the precentral gyrus (henceforth, PrC). The other regions were in the planum temporale (PT), superior temporal gyrus (STG) anterior to Heschl's gyrus, pars opercularis (POp), and pars triangularis (PTr) extending into the left anterior insula (Ins) (for details, see Fig. 3 and Table 3). Activation in these regions was higher for all bilingual groups, even when the Greeks were responding in their native language (see Fig. 3). Moreover, activation did not differ for Groups 2a, 2b, and 2c, or between the first versus second language in the Greek-English bilinguals (Group 2c), consistent with the high proficiency of all our participants (Perani et al. 1998; Chee et al. 2001; Meschyan and Hernandez 2006).

When we lowered the statistical threshold ($P < 0.001$ uncorrected), the right homologues of STG, POp, and PTr also

Table 3

Areas showing increased activation for bilinguals compared with monolinguals

Anatomical region	MNI coordinates			Z scores for					
	x	y	z	Analysis 1			Analysis 2		
				N/R	Sem	Int	N/R	Art	Int
PT Planum temporale	-56	-30	+14	5.8	ns	5.5	5.4	3.9	ns
PrC Dorsal precentral gyrus	-48	-16	+42	5.3	ns	4.6	5.1	4.1	ns
STG Superior temporal gyrus	-60	-10	+2	4.7	ns	4.5	4.2	3.1	ns
POp Pars opercularis	-56	+10	+18	5.3	ns	4.1	5.4	3.0	3.2
PTr Ventral pars triangularis	-42	+26	+6	5.3	ns	4.8	5.1	ns	3.1
Ins Anterior insula	-36	+14	+4	4.6	ns	4.4	4.5	ns	3.1

Note: N/R, naming and reading aloud; Sem, semantic decisions; Art, articulating "1,2,3"; Int, interaction; ns, not significant.

showed greater activation in bilinguals than in monolinguals during the naming and reading tasks (MNI coordinates = [+56, -8, +2], [+56, +4, +14], and [+42, +22, +2]). However, there was no evidence for increased activation in the left head of caudate or anterior cingulate cortex that have previously been associated with language control (Crinion et al. 2006; Khateb et al. 2007; Abutalebi et al. 2008). The most likely explanation is that our paradigm did not involve any language switching whereas previous studies have reported activation in these regions when bilinguals knowingly switch between one language and another.

The effect of bilingualism did not differ for picture naming or reading aloud (whole brain analysis or region of interest analysis), and there were no areas that were more activated for monolinguals than bilinguals. There were no significant group differences during semantic decisions on words or pictures ($P > 0.001$ uncorrected), and the interaction of group by task confirmed that the effect of bilingualism was greater on naming and reading than semantic decision in all the regions identified above (see Table 3).

Analysis 2: Activation for Bilinguals versus Monolinguals for Naming and Reading Aloud versus Articulation

The effect of bilingualism during naming and reading aloud was highly significant in the same 5 regions identified in Analysis 1. In PT, PrC, STG, there was also an effect of bilingualism for articulating 1,2,3, with no significant difference in the effect of bilingualism during naming, reading aloud, or articulating 1,2,3. In contrast, the effect of bilingualism in POp and PTr/Ins was higher for naming and reading aloud than articulating 1,2,3 (for details, see Table 3). There were no other group effects.

Analysis 3: Correlating the Ability to Control Verbal Interference with Activation in Areas Associated with Bilingualism

In monolinguals, we found a negative relationship between CR_{RT} and activation for picture naming and reading aloud in POp ($Z = -3.7$, $P < 0.001$) and PTr ($Z = -2.5$, $P < 0.01$). This indicates more activation with less conflict, consistent with a role for POp and PTr in the control of interference during monolingual picture naming and reading. In contrast, there were no significant correlations between CR_{RT} and activation in the bilinguals for either the native language CR_{RT} or their nonnative language CR_{RT} and no correlations between CR_{RT} and activation were observed in PT, STG, or PrC, in either the monolinguals or the bilinguals (native or nonnative language) even when the statistical threshold was reduced to $P < 0.05$

uncorrected. Conceivably, the absence of a correlation between CR_{RT} and bilingual activation is because the control of interference was consistently greater for bilinguals than monolinguals during the naming and reading aloud tasks.

The results of Analysis 3 show that, although POP and PTr activation for monolingual naming and reading aloud was not significant at the group level (see right side of Fig. 3), activation increased in these areas for monolingual individuals who were better at controlling verbal interference in the Stroop task. This suggests that POP and PTr are involved in controlling interference from competing words in the same language. Greater activation in POP and PTr for bilingual than for

monolingual naming and reading is therefore consistent with our hypothesis that controlling interference is more demanding for bilinguals because words in the other language also compete for selection.

Analysis 4: Proficiency and Age of Acquisition

Although our selection criteria only included bilinguals who were highly proficient in English, there was variance across the sample. Analysis 4 therefore tested whether activation in the bilinguals varied either with written word knowledge (as measured by accuracy and RT on the lexical decision task from the PALPA) or with age of acquisition. Consistent with our

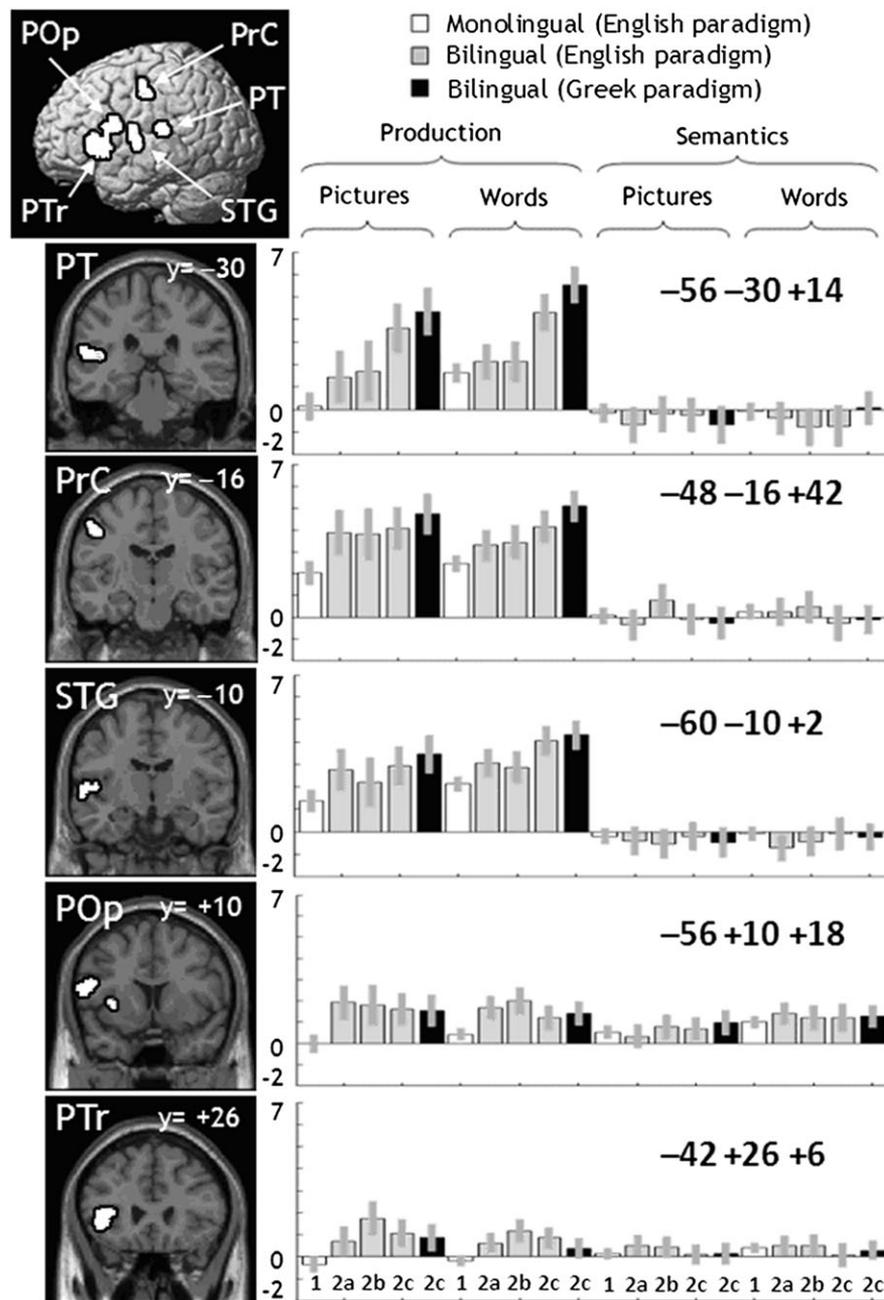


Figure 3. Main effect of Bilingualism. This is a visualization of Table 3. A sagittal rendering and coronal sections are given on the left. F-maps are on the right, with groups and conditions along the x-axis, and values along the y-axis (within the range $[-2, 7]$). Conditions are arranged into tasks (naming pictures, reading words aloud, matching semantically related pictures, matching alphabetic strings). Conditions are also colored white for monolinguals, gray for bilinguals in the English paradigm, and black for bilinguals in the Greek paradigm.

selection aims, there were no significant correlations for picture naming and reading aloud ($P < 0.05$ uncorrected) in any of the regions we identified as more activated for bilinguals than for monolinguals. At the whole brain level, we observed a negative correlation between word knowledge and activation in the right cerebellum [$+32, -82, -24$] ($Z = -4.8$). While this is consistent with prior studies showing more cerebellar activation with lower proficiency (Liu et al. 2010), it is not discussed further here as our focus was on differences between bilinguals and monolinguals rather than the effect of proficiency within bilinguals.

Analysis 5: The Effect of the Number of Languages Spoken

Many of our participants spoke more than 2 languages (for details, see Table 1). Therefore, Analysis 5 tested whether activation in the regions where we found differences in bilinguals and monolinguals above varied with the number of languages spoken. Three groups of participants were included. Group 1 ($n = 36$) was the same as Group 1 in Analyses 1 and 2 (i.e., all the monolinguals); Group 2 ($n = 9$) contained the bilinguals who spoke no more than 2 languages (irrespective of whether they were in Groups 2a, 2b, or 2c in Analyses 1 and 2); and Group 3 ($n = 21$) contained bilinguals who spoke 3 or more languages. For each group, we included the contrast images for picture naming—fixation and reading aloud—fixation. Therefore, there were a total of 6 different conditions.

The comparison of naming and reading activation for speaking 1, 2, or 3+ languages confirmed that activation in all 5 of our regions of interest (i.e., PT, PrC, STG, POp, PTr/Ins) was greater for those who spoke 2 languages or 3+ languages than for those who spoke 1 language (see Fig. 4). In addition, this analysis found that activation in PT and PTr/Ins was greater ($P < 0.001$) for those who spoke 3+ languages than for those who spoke 2 languages. Further work may explore this intriguing finding. However, it suffices for now to demonstrate that the effect of bilingualism we report in this paper pertained to participants who spoke 2 languages as well as those who spoke more than 2 languages.

Discussion

This study shows that, compared with monolinguals, bilinguals have increased activation in 5 left frontal and temporal regions when they name pictures or read words aloud in a single language context, either in their native language or in their nonnative language. By considering the location of these effects, and comparing them with others reported in the literature, we show where, when, and why brain activation differs in bilinguals and monolinguals during simple speech production tasks.

Where Did Activation Increase for Bilinguals?

In the whole brain search, 5 regions showed higher activation for bilinguals than monolinguals during naming and reading aloud versus either fixation or semantic decisions; these were labeled PT, PrC, STG, POp, PTr/Ins (see Fig. 3). With the exception of the left STG, all these “bilingual” areas have previously been associated with frequency effects within the native language (Graves et al. 2007; see Table 2) and POp and PTr have also been associated with the control of interference in bilingual studies (Rodriguez-Fornells et al. 2002, 2005). The

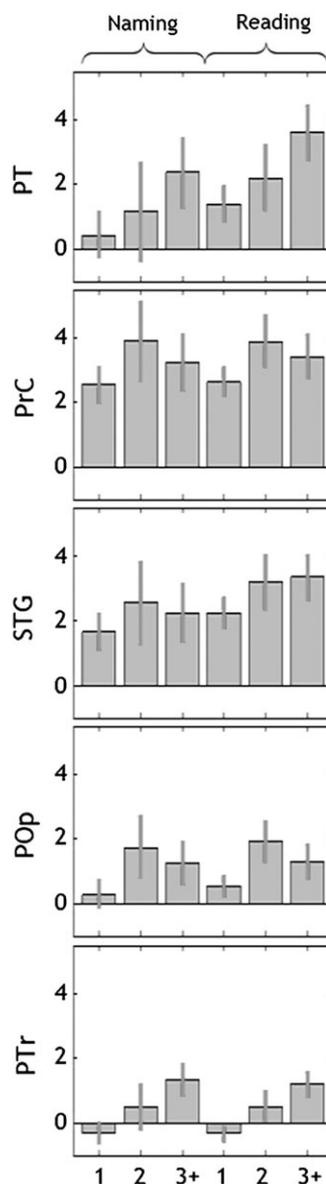


Figure 4. Effect of number of languages (Analysis 5). Parameter estimates for number of languages in 5 significant regions (from the whole brain analysis: PT, PrC, STG, POp, PTr/Ins). The x-axes are number of languages (1 = monolingual, 2 = exactly bilingual, 3+ = more than exactly bilingual). The y-axes are the estimated effect sizes; the variance is standard error.

only area where we predicted but did not detect activation was the left middle frontal area [$-40, +36, +32$], which Rodriguez-Fornells et al. (2005) associated with phonological interference during a task that required participants to decide if the first letter of an object’s name was a German vowel. The left middle frontal activation reported by Rodriguez-Fornells et al. (2005) may therefore be related to the vowel discrimination decision that took place after retrieving picture names in their experiment.

Our observation that the effect of bilingualism was located in left frontal and temporal regions where an effect of frequency has been observed during picture naming (Graves et al. 2007) suggests that bilingualism places additional demands on processing that is also involved in monolingual speech production. This

explains an established observation in the bilingual literature, which is that bilinguals seem to incur an extra processing cost, for example, when naming pictures (Gollan et al. 2005). The correspondence between the effect of bilingualism and the control of lexical interference in POp and PTr further suggests that the effects of bilingualism and frequency might be partly explained by the control of interference. However, one region—left STG—has not previously been associated with either low- versus high-frequency picture naming or interference (though see Maess et al. 2002). Without further consideration, it might be concluded that the STG was involved in processing that was specific to bilinguals, consistent with the possibility that the control of competition between 2 languages involves different processing resources to the control of competition within a single language. This hypothesis is not, however, supported when we consider the response properties of this region below.

When Did Activation Increase for Bilinguals?

Here, we consider when in the multistage processing stream the effect of bilingualism arose. First, we established that the differences between bilinguals and monolinguals arose when the names of the stimuli needed to be retrieved and articulated. This inference was based on finding that 1) there were no group differences during semantic decisions on either pictures or words; 2) activation increased for bilinguals relative to monolinguals when the task was naming and reading aloud; and 3) these group differences in naming and reading aloud were observed irrespective of whether the baseline condition was fixation or semantic decisions.

Second, we divided the regional effects according to whether they arose at the level of word retrieval or articulation. This involved comparing the effect of bilingualism for 3 different speech production tasks: naming, reading aloud, and articulating 1,2,3. We found that in 3 of our 5 regions of interest (i.e., PrC, PT, STG), group differences were emerging at the level of articulation or postarticulatory processing (e.g., auditory-motor feedback; Dhanjal et al. 2008) because they were observed for all 3 tasks with no interaction between group and task. By contrast, the effect of bilingualism in the other 2 areas (POp and PTr/Ins) was greater during naming and reading aloud than during articulation ($P < 0.001$), consistent with a role in word retrieval and articulatory planning rather than articulation per se. Interestingly, the effect of bilingualism on brain activation was not greater for naming than reading aloud. At one level this might be surprising because word retrieval is facilitated during reading by sublexical associations between visual parts (letters) and phonology. On the other hand, the use of sublexical associations may be suppressed in bilinguals, particularly in languages like German and Italian that use the same orthography but with different phonological associations (Nosarti et al. 2010).

Third, for each of the areas showing group differences in activation, we examined the effect of task, within group. This associated PT, PrC, and STG to articulation and/or postarticulatory processing because these areas were more activated by naming and reading aloud than semantic decisions with no difference in activation during naming, reading, or articulating 1,2,3 (Analyses 1 and 2). In contrast, PTr and Pop were associated with the ability to control verbal interference within language (Analysis 3).

Fourth, we consider how the areas showing an effect of bilingualism in our study had been interpreted in prior studies (for a review, see Price 2010). We note that previous studies have associated PTr with word retrieval (Devlin et al. 2003) and cognitive control (Koechlin et al. 2003); the adjacent insula (Ins) with articulatory planning (Dronkers 1996; Wise et al. 1999; Brown et al. 2009; Fridriksson et al. 2009; Moser et al. 2009; Shuster 2009); POp with controlling word retrieval and articulatory sequences, PrC with motor output of articulatory sequences (e.g., Fridriksson et al. 2009); and PT and STG with the postarticulatory processing of speech (see Price 2010).

The association of PT and STG with articulation may seem surprising. However, the STG region that we see activated in all articulation tasks is one that responds to prelexical phonemic processing of auditory input during perception (Britton et al. 2009; Leff et al. 2009). Although none of our conditions involved external auditory input, the output from articulation is itself an auditory response that results in auditory activation (Price et al. 1996; Bohland and Guenther 2006). Indeed, the coordinates of our STG activation [MNI: -60, -10, +2] are remarkably similar to the lateral STG activation [MNI: -56, -12, 0] associated with auditory processing during speech output in Dhanjal et al. (2008). Likewise, activation in our PT area [MNI: -56, -30, +14] is similar to the activation that Takaso et al. (2010) report [MNI: -60, -30, +8] when participants were given delayed feedback of their own voices over a pair of headphones (compared with nondelayed feedback). This is consistent with proposals that STG and PT are involved in auditory-motor feedback during speech production (Dhanjal et al. 2008; Zheng et al. 2009) and that increased activation in nonnative speech processing may reflect the successful use of articulatory-auditory and articulatory-orosensory feedback (Callan et al. 2004).

In brief, the functional response properties of the areas associated with bilinguals suggest that bilinguals increase activation relative to monolinguals in areas involved in controlling verbal interference during word retrieval and articulatory planning (PTr/Ins), control of articulatory sequences (POp), articulation (PrC), auditory processing of speech output (STG), and auditory-motor feedback (PT).

Why Does Activation Increase for Bilinguals?

Why might the demands on the various processes described above be relatively greater in bilinguals as compared with monolinguals? We will evaluate our hypotheses in light of the functional properties associated with each region.

Consider first the response in PTr. Our data show that within monolingual speakers, PTr acts to control interference in order to ensure correct word selection. The frequency effect in PTr, that has been reported in monolinguals (Graves et al. 2007), can then be understood in terms of the demands on interference control, consistent with behavioral data showing that lexical selection is affected by word frequency (Navarrete et al. 2006; Kittredge et al. 2008). The demand on this region is greater in bilinguals because of the additional need to control interference from words that are not in the target language (see e.g., Rodriguez-Fornells et al. 2002; Abutalebi and Green 2007; Abutalebi 2008).

In POp, activation has been associated with predicting semantic and articulatory sequences and the control of verbal interference in monolinguals (see above). The need for such predictions may vary with the level of competition from words

with similar sounds and meanings. As in PTR, the data point toward a domain general system that is used in both monolinguals and bilinguals but is more activated in bilinguals because of competition between languages as well as within languages. In contrast, activation in the PrC area has been associated with articulation but not the control of interference. In bilinguals, articulatory activation in PrC might be higher because the motor plan and execution of each word is less rehearsed, as predicted by the reduced frequency hypothesis. Finally, activation in STG and PT has been associated with auditory-motor feedback from the articulated response. Plausibly, this is greater in bilinguals than monolinguals because bilinguals may monitor the spoken response more carefully to ensure that it is articulated with the accent, intonation, and other features associated with the target language.

In sum, the effect of bilingualism on regional activations in naming and reading is best explained by greater demands on the processes of word retrieval, articulation, and postarticulatory monitoring that are in common with word processing in monolinguals.

Summary and Conclusions

In this study, we have shown increased activation for bilinguals relative to monolinguals during overt picture naming and reading aloud, even when bilinguals are only responding in their native language. The areas where these effects were observed are remarkably consistent with those previously associated with low- versus high-frequency picture naming in one's native language and the control of interference in bilinguals as they respond in a dual language context. Our findings suggest that bilinguals increase processing within a system that is also used in monolinguals (for a similar rationale, see Abutalebi and Green 2007). However, they contrast sharply with the idea of a unique and advantageous bilingual system that exploits resources that are untapped in monolinguals (Kovelman, Baker, et al. 2008; Kovelman, Shalinsky, et al. 2008). By including multiple tasks, we have been able to interpret the function of the areas where activation is higher in bilinguals than monolinguals. By including multiple groups, and only testing in a single language context, we were also able to control for differences between native versus nonnative language. Finally, our study goes beyond an exploration of where effects arise in bilinguals, and specifically tests evidence for competing interpretations of the data. Our conclusions offer novel insight into the effect of bilingualism on brain function and emphasize that the advantage of being bilingual comes at the expense of increased demands on word retrieval and articulation, even in simple picture naming and reading tasks.

Funding

Wellcome Trust.

Notes

We would like to thank our participants for their time as well as the radiographers who helped us acquire the data. *Conflict of Interest*: None declared.

References

Abutalebi J. 2008. Neural aspects of second language representation and language control. *Acta Psychol.* 128:466-478.

- 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, Green DW. 2007. Bilingual language production: the neurocognition of language representation and control. *J Neuro-linguist.* 20:242-275.
- Andersson JLR, Hutton C, Ashburner J, Turner R, Friston K. 2001. Modelling geometric deformations in EPI time series. *Neuroimage.* 13:903-919.
- Bohland JW, Guenther FH. 2006. An fMRI investigation of syllable sequence production. *Neuroimage.* 32:821-841.
- Britton B, Blumstein SE, Myers EB, Grindrod C. 2009. The role of the spectral and durational properties on hemispheric asymmetries in vowel perception. *Neuropsychologia.* 47:1096-1106.
- Brown S, Laird AR, Pfordresher PQ, Thelen SM, Turkeltaub P, Liotti M. 2009. The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain Cogn.* 70:31-41.
- Callan DE, Jones JA, Callan AM, Akahane-Yamada R. 2004. Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *Neuroimage.* 22:1182-1194.
- Chee MW, Hon N, Lee HL, Soon CS. 2001. Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments: blood oxygen level dependent. *Neuroimage.* 13:1155-1163.
- 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.
- Devlin JT, Matthews PM, Rushworth MF. 2003. Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *J Cogn Neurosci.* 15:71-84.
- Dhanjal NS, Handunnetthi L, Patel MC, Wise RJ. 2008. Perceptual systems controlling speech production. *J Neurosci.* 28:9969-9975.
- Dronkers NF. 1996. A new brain region for coordinating speech articulation. *Nature.* 384:159-161.
- Fridriksson J, Moser D, Ryalls J, Bonilha L, Rorden C, Baylis G. 2009. Modulation of frontal lobe speech areas associated with the production and perception of speech movements. *J Speech Lang Hear Res.* 52:812-819.
- Gollan TH, Acenas LA. 2004. What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *J Exp Psychol Learn Mem Cogn.* 30:246-269.
- Gollan TH, Montoya RI, Bonanni MP. 2005. Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology.* 19:278-287.
- Gollan TH, Montoya RI, Cera CM, Sandoval TC. 2008. More use almost always means a smaller frequency effect: aging, bilingualism, and the weaker links hypothesis. *J Mem Lang.* 58:787-814.
- Gollan TH, Montoya RI, Werner GA. 2002. Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology.* 16:562-576.
- Graves WW, Grabowski TJ, Mehta S, Gordon JK. 2007. A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. *J Cogn Neurosci.* 19:617-631.
- Green DW. 1986. Control, activation, and resource: a framework and a model for the control of speech in bilinguals. *Brain Lang.* 27:210-223.
- Green DW. 1998. Mental control of the bilingual lexico-semantic system. *Biling Lang Cogn.* 1:67-81.
- Green DW, Grogan A, Crinion J, Ali N, Sutton C, Price CJ. 2010. Language control and parallel recovery of language in individuals with aphasia. *Aphasiology.* 24:188-209.
- Grogan A, Green DW, Ali N, Crinion J, Price CJ. 2009. Structural correlates of semantic and phonemic fluency ability first and second languages. *Cereb Cortex.* 19:2690-2698.
- Grosjean F. 1992. Another view of bilingualism. In: Harris RJ, editor. *Cognitive processing in bilinguals.* Amsterdam: Elsevier. p. 51-62.
- Grosjean F. 1998. Studying bilinguals: methodological and conceptual issues. *Biling Lang Cogn.* 1:131-140.

- Grosjean F. 2001. The bilingual's language modes. In: Nicol J, editor. *One mind, two languages: bilingual language processing*. Oxford: Blackwell. p. 1-22.
- Hu W, Lee HL, Zhang Q, Liu T, Geng LB, Seghier ML, Shakeshaft C, Twomey T, Green DW, Yang YM, et al. 2010. Developmental dyslexia in Chinese and English populations: dissociating the effect of dyslexia from language differences. *Brain*. 133:1694-1706.
- Ivanova I, Costa A. 2008. Does bilingualism hamper lexical access in speech production. *Acta Psychol*. 127:277-288.
- Jescheniak JD, Schriefers H. 1998. Discrete serial versus cascaded processing in lexical access in speech production: further evidence from the coactivation of near-synonyms. *J Exp Psychol Learn Mem Cogn*. 24:1256-1274.
- Kay J, Lesser R, Coltheart M. 1992. *Psycholinguistic assessment of language processing in aphasia*. Hove (UK): Psychology Press.
- 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.
- Kittredge AK, Dell GS, Verkuilen J, Schwartz MF. 2008. Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cogn Neuropsychol*. 25:463-492.
- Koechlin E, Ody C, Kouneiher F. 2003. The architecture of cognitive control in the human prefrontal cortex. *Science*. 302:1181-1185.
- Kovelman I, Baker SA, Petitto LA. 2008. Bilingual and monolingual brains compared: a functional magnetic resonance imaging investigation of syntactic processing and a possible "neural signature" of bilingualism. *J Cogn Neurosci*. 20:153-169.
- Kovelman I, Shalinsky MH, Berens MS, Petitto LA. 2008. Shining new light on the brain's "bilingual signature": a functional Near Infrared Spectroscopy investigation of semantic processing. *Neuroimage*. 39:1457-1471.
- Kovelman I, Shalinsky MH, White KS, Schmitt SN, Berens MS, Paymer N, Petitto LA. 2009. Dual language use in sign-speech bimodal bilinguals: fNIRS brain-imaging evidence. *Brain Lang*. 109:112-123.
- Leff AP, Iverson P, Schofield TM, Kilner JM, Crinion JT, Friston K, Price CJ. 2009. Vowel specific mismatch responses in the anterior superior temporal gyrus: an fMRI study. *Cortex*. 45:517-526.
- Liu H, Hu Z, Guo T, Peng D. 2010. Speaking words in two languages with one brain: neural overlap and dissociation. *Brain Res*. 1316:75-82.
- Long DL, Prat CS. 2002. Working memory and Stroop interference: an individual differences investigation. *Mem Cognit*. 30:294-301.
- Maess B, Friederici AD, Damian M, Meyer AS, Levelt JM. 2002. Semantic category interference in overt picture naming: sharpening current density localization by PCA. *J Cogn Neurosci*. 14:455-462.
- Mägiste E. 1979. The competing language systems of the multilingual: a developmental study of decoding and encoding processes. *J Verb Learn Verb Behav*. 18:79-89.
- Meschyan G, Hernandez AE. 2006. Impact of language proficiency and orthographic transparency on bilingual word reading: an fMRI investigation. *Neuroimage*. 29:1135-1140.
- Moser D, Fridriksson J, Bonilha L, Healy EW, Baylis G, Baker JM, Rorden C. 2009. Neural recruitment for the production of native and novel speech sounds. *Neuroimage*. 46:549-557.
- Navarrete E, Basagni B, Alario XF, Costa A. 2006. Does word frequency affect lexical selection in speech production? *Q J Exp Psychol*. 10:1681-1690.
- Nosarti C, Mechelli A, Green DW, Price C. 2010. The impact of second language learning on semantic and nonsemantic first language reading. *Cerebral Cortex*. 20:315-327.
- Papoutsis M, de Zwart JA, Jansam JM, Pickering MJ, Bednar JA, Horwitz B. 2009. From phonemes to articulatory codes: an fMRI study of the role of Broca's area in speech production. *Cereb Cortex*. 19:2156-2165.
- Perani D, Paulesu E, Galles NS, Dupoux E, Dehaene S, Bettinardi V, Cappa SF, Fazio F, Mehler J. 1998. The bilingual brain: proficiency and age of acquisition of the second language. *Brain Lang*. 121:1841-1852.
- Peterson RR, Savoy P. 1998. Lexical selection and phonological encoding during language production: evidence for cascaded processing. *J Exp Psychol Learn Mem Cogn*. 24:539-557.
- Price CJ. 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann N Y Acad Sci*. 1191:62-88.
- Price CJ, Wise RJ, Warburton EA, Moore CJ, Howard D, Patterson K, Frackowiak R, Friston K. 1996. Hearing and saying: the functional neuro-anatomy of auditory word processing. *Brain*. 119:919-931.
- Pyers JE, Gollan TH, Emmorey K. 2009. Bimodal bilinguals reveal the source of tip-of-the-tongue states. *Cognition*. 112:323-329.
- Ransdell SE, Fischler I. 1987. Imagery skill and preference in bilinguals. *Appl Cogn Psychol*. 5:97-112.
- Rodriguez-Fornells A, Rotte M, Heinze HJ, Nössel T, Münte TF. 2002. Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*. 415:1026-1029.
- Rodriguez-Fornells A, van der Lugt A, Rotte M, Britti B, Heinze HJ, Münte TF. 2005. Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *J Cogn Neurosci*. 17:422-433.
- Shuster LI. 2009. The effect of sublexical and lexical frequency on speech production: an fMRI investigation. *Brain Lang*. 111:66-72.
- Stroop JR. 1935. Studies of interference in serial verbal reactions. *J Exp Psychol*. 18:643-662.
- Takaso H, Eisner F, Wise RJ, Scott SK. 2010. The effect of delayed auditory feedback on activity in the temporal lobe while speaking: a positron emission tomography study. *J Speech Lang Hear Res*. 53:226-236.
- van Heuven WJ, Dijkstra T. 2010. Language comprehension in the bilingual brain: fMRI and ERP support psycholinguistic models. *Brain Res Rev*. 64:104-122.
- Veltman DJ, Mechelli A, Friston K, Price CJ. 2002. The importance of distributed sampling in blocked functional magnetic resonance imaging designs. *Neuroimage*. 17:1203-1206.
- Wise RJ, Greene J, Büchel C, Scott SK. 1999. Brain regions involved in articulation. *Lancet*. 353:1057-1061.
- Wu YJ, Thierry G. 2010. Chinese-English bilinguals reading English hear Chinese. *J Neurosci*. 30:7646-7651.
- Zheng ZZ, Munhall KG, Johnsrude IS. 2009. Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *J Cogn Neurosci*. 22:1770-1781.