



Neural activity during Stroop colour-word task performance in late proficient bilinguals: a functional Magnetic Resonance Imaging study

Karen E. Waldie¹, Gjurgjica Badzakova-Trajkov¹, Branka Miliivojevic² and Ian J. Kirk¹

¹ University of Auckland, New Zealand

² Utrecht University, The Netherlands

Abstract

The aim of this study was to identify which neural substrates are engaged during manual Stroop task performance and compare the activation between 8 late proficient Macedonian-English bilinguals and 10 matched English monolinguals. During functional magnetic resonance imaging (fMRI) adult participants performed four Stroop task conditions in a block design (control, congruent, semantically incongruent, response incongruent). Here, we focussed on differences in activation between the two groups in two contrasts: (1) the presence of task-irrelevant information that conflicts at both the response and semantic level (response incongruent versus congruent); and (2) the presence of competing task-relevant information that conflicts at the response level only (response incongruent versus semantically incongruent). For the first contrast, comparisons between monolinguals (English) and the first language of bilinguals (Macedonian) revealed greater activation in the posterior regions of monolinguals, possibly indicative of a more elaborate stimulus evaluation process. For the second contrast, bilinguals showed less activation in anterior cingulate and prefrontal regions than monolinguals. As poor performance on the Stroop task has been found to be associated with greater activation in these two regions, we interpret this in light of earlier ideas regarding cognitive control and the advantages of bilingualism. That is, less effortful word-form processing might indicate that bilinguals are more effectively dealing with response conflict. Together the current findings suggest that learning two languages may modify the neural substrates engaged during tasks that rely on executive functioning. **Keywords:** bilingualism, fMRI, interference, conflict, anterior cingulate, DLPFC; cognitive neuroscience

Received 25 September 2009; received in revised form 6 December 2009; accepted 6 December 2009. Available on line 29 December 2009.

Introduction

The effect that bilingualism has on brain development and functioning, and the possibility that there is a cognitive advantage to bilingualism are growing concerns in both linguistic and cognitive neuroscience research. It has been recently shown, for example, that bilingualism may have a positive effect on executive functioning and acts as a protective factor against decline in cognitive functioning with age (Bialystok, 2009). According to a number of researchers, the positive effects observed in the bilinguals' cognitive processing, in terms of better selective attention or inhibitory control, are due

to having some form of monitoring system when using two languages on a consistent basis in order to prevent intrusions from the non-target language (e.g., Bialystok, 2001; Costa, Albareda, & Santesteban, 2008; Green, 1998). Indeed, the need for constant monitoring is evident from behavioural research showing that during lexical access both languages are engaged at the same time (Costa, Caramazza, & Sebastian-Galles, 2000; Marian, Spivey, & Hisch, 2003).

The effect of bilingualism on executive processes has been typically investigated with the Simon task. In this task, participants are instructed to press a right response key for a blue square presented on the screen and a left response key for a red square. The stimuli are then presented on either the correct side of the screen (congruent position) or the incorrect side of the screen (incongruent position). A bilingual advantage, in terms of faster reaction times and/or smaller Simon effect (incongruent response times minus congruent response times), have been reported in both bilingual children (Martin-Rhee & Bialystok, 2008) and bilingual adults (Bialystok, Craik, Klein, & Viswanathan, 2004).

Karen E. Waldie, Gjurgjica Badzakova-Trajkov and Ian J. Kirk, Research Centre for Cognitive Neuroscience, Department of Psychology, University of Auckland, New Zealand. Branka Miliivojevic, Department of Experimental and Developmental Psychology, Utrecht University, The Netherlands. Correspondence regarding this article should be directed to: Karen E. Waldie, Department of Psychology, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand, phone: +64 9-3737599 ext.88521. E-mail: k.waldie@auckland.ac.nz.

A similar bilingual advantage has been observed on a range of different tasks thought to depend on executive control. For example, Bialystok and Shapero (2005) showed that bilingual children perform better than monolingual children when reversing ambiguous figures (images that can have alternate interpretations like the vase-face image). In this task one needs to selectively attend to certain aspects of the stimuli and resolve conflict from competing alternatives. Bialystok, Craik and Ryan (2006) further found that bilinguals had a performance advantage in an antisaccade task where participants had to refrain from looking at the stimulus cue direction and to make an antisaccadic movement in the opposite direction. Finally, Carlson and Meltzoff (2008) employed a number of different tasks with bilingual children, including the Simon task, and found that they were better than monolinguals on tasks that required the participant to manage conflicting attentional demands.

Despite the growing list of reports that found a bilingual advantage on so-called “executive” or conflict tasks, we know little about which neural substrates may give rise to this behavioural advantage. Executive processes are generally associated with the prefrontal cortex, in particular the dorsolateral prefrontal cortex (DLPFC; Petrides, 2000) and anterior cingulate cortex (ACC; Cohen, Botvinick, & Carter 2000). Parietal regions, such as the inferior parietal cortex and supramarginal gyrus, have also been found to be involved (Wager, Jonides, & Reading, 2004).

The importance of these frontal and parietal regions in executive processes is also evident from neuropsychological work with frontal lesion patients. These patients often perform poorly on tasks that rely on intact executive functioning, such as the Wisconsin card sorting test, phonemic verbal fluency, and the Stroop colour-word interference task (Alvarez & Emory, 2006). However, the exact roles these brain regions play in different executive processes are still under investigation. In general, ventrolateral prefrontal regions are concerned with simple short-term operations, whereas middorsal prefrontal regions are thought to carry out higher-level executive operations, such as monitoring (Petrides, 1995). More specifically, the DLPFC is thought to exert top-down control by maintaining task relevant information about goals or task rules, and uses this information to guide processing in posterior cortical regions (Miller & Cohen, 2001). The ACC has been associated with error detection and conflict resolution (e.g., van Veen & Carter, 2002).

A very small number of studies have used neuroimaging techniques to investigate the effects of bilingualism on executive functioning at the neural substrate level. In a magneto-encephalography (MEG) study, Bialystok et al. (2005) used the Simon task with early proficient bilingual adults (French-English and Cantonese-English) and English monolinguals and

found that fast responding in bilinguals was associated with activation in cingulate, superior frontal, and inferior frontal regions in the left hemisphere. Fast responding in monolinguals was associated with left middle frontal activation. In an fMRI study, Kovelman (2006) used a nonverbal cognitive task where Spanish-English bilinguals and English monolinguals attended to the position or direction of an arrow that appeared on the screen. In this study, bilinguals showed greater activation in the left inferior parietal region and left supramarginal gyrus compared to monolinguals. Taken together, both studies showed differences in activation between bilinguals and monolinguals during non-linguistic cognitive tasks that required some type of inhibitory control. This suggests that there is differential engagement of brain areas supporting frontal executive functions in bilinguals compared to monolinguals.

The current study employed the Stroop colour-word task, considered one of the classical measures of executive functioning (Lezak, 1995), during concurrent fMRI to explore the following hypothesis: Does the constant management of the two language systems in bilinguals (i.e., language switching) promote positive changes in executive functions?

The Stroop effect occurs when participants are asked to name the colour of a word that is incongruent with the required colour response (e.g., RED printed in blue ink) (Stroop, 1935). This typically results in increased response time relative to both the congruent condition (e.g., RED printed in red ink) and the control condition (e.g., JUMP printed in red ink). Unlike the non-linguistic tasks described above in earlier studies, the Stroop task is contingent on printed words and reading. This enabled us to investigate the effect of bilingualism on cognitive control processes for each of the two languages.

Although involvement of the ACC and DLPFC in Stroop task performance has been readily observed in neuroimaging research with monolingual adults (MacLeod & MacDonald 2000), disagreement exists as to the specific roles they play (Cohen et al., 2000). For example, Posner and Dehaene (1994) argued that the ACC is involved in the top-down implementation of selection and processing of stimuli that need to be acted upon and thus has an ‘executive’ role in the control of attention. This is further supported by studies showing ACC activation in conditions where there is no conflict (e.g., Roelofs, van Turrenout, & Coles, 2006).

Others, however, have viewed the ACC as an error detector that monitors performance (Carter, Braver, Barch, Botvinick, Noll, & Cohen, 1998; MacDonald, Cohen, Stenger, & Carter, 2000) or detects the presence of conflicting information (van Veen & Carter, 2002), thus supporting the conflict monitoring theory of ACC function (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Carter et al., 1998). According to this view, the ACC will be activated in conditions where there is

conflicting information, as it is recruited to firstly detect such events and secondly to signal the need for attentional control. Milham and Banich (2005) suggested that top-down attentional control itself may in fact be mediated by the DLPFC (e.g., MacDonald et al., 2000). It is clear, however, that these regions are almost always involved in situations where the possibility of a conflict is increased.

In the present study late and proficient bilinguals and a matched monolingual group were tested using a paradigm that is slightly different to the traditional Stroop task, as incongruent stimuli were further divided into semantically incongruent stimuli and response incongruent stimuli. There was a two-choice button-press with two colours associated with each button/finger (van Veen & Carter, 2005). Semantically incongruent stimuli contained semantic conflict information only, as the response key was the same for both the target colour and the colour word. Response incongruent stimuli contained both semantic conflict and response conflict information, as the response key for the target colour was different from the colour word.

This allowed for comparisons both between groups and between languages on two cognitive components of the Stroop task: (1) the presence of task-irrelevant information that conflicts at both the response and semantic level (interference) (response incongruent condition versus congruent condition) which is the most commonly evaluated contrast; and (2) the presence of competing task-relevant information that conflicts at the response level (response conflict) (response incongruent condition versus semantically incongruent condition) which is particularly important for the conflict resolution process. The latter contrast allowed us to exclude general semantic processing differences between the groups (and languages) from confounding the conflict resolution process.

It was of particular interest to identify the neural areas that are recruited during Stroop task performance and to determine whether these areas differ as a function of our predefined contrasts. Overall, for both contrasts, we expected that activation in anterior cingulate and prefrontal regions may be negatively associated with task performance in bilinguals.

Method

Participants

English monolingual participants consisted of 5 males and 5 females (Mean age = 28.4, $SD = 7.4$ years), all of whom were undergraduate or postgraduate students attending The University of Auckland, New Zealand.

Macedonian-English bilingual participants were 5 males and 3 females (Mean age = 24.87, $SD = 2.7$ years) from the Macedonian community in Auckland. All were well educated and either university students or professionals. The Macedonian language is related to English in that they are both from the Indo-European

group of languages. Macedonian is from the Slavic Indo-European language group, whereas English is from the Germanic Indo-European group.

All participants were right handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Independent samples t-tests revealed no differences between the groups in age or handedness. Participants had no history of neurological disorder and all had normal or corrected to normal vision. Ethics approval for this project was given by The University of Auckland Human Participants Ethics Committee and written consent was obtained from all participants.

Bilingual participants were all considered "late" bilinguals in that they learned English as a second language (L2) after the age of 7 years (between 9-17 years of age; mean age = 11.25, $SD = 2.55$). English proficiency was assessed using a computerised Quick Placement Test (QPT, 2001) which is a 15-20 minute adaptive computer-based test (computer selects difficulty of questions based on previous answers) and measures listening, reading, grammar and vocabulary. The test has been validated in 20 countries by more than 5,000 students of English. Participants are categorized into one of six levels based on their score (out of 100): 0-Beginner (0-39); 1-Elementary (40-49); 2-Lower Intermediate (50-59); 3-Upper Intermediate (60-69); 4-Advanced (70-79); and 5-Very Advanced (80-100).

The mean English proficiency score for the bilinguals was 84.75 ($SD = 6.30$). The majority of these participants reported using both languages equally in everyday life in New Zealand, and about 37% reported using English more than their native Macedonian language (L1). Contexts of use for L1 were mainly at home (100%), with friends and social outings (37.5%), whereas contexts of use for L2 were mainly work (87.5%), shopping (62.5%), and social outings (25%). The time spent in the L2 environment for the bilingual group ranged from 6-11 years, with a mean of 8.88. This indicates that the bilingual participants had spent a significant portion of their life in the L2 environment given their age.

Stimuli

Stimuli were presented in Courier New bold, 35 font size. Stimuli were presented in red, blue, green or yellow. Red and blue target colours were mapped onto the left mouse button whereas green and yellow were mapped onto the right mouse button.

Stimuli in the control condition were neutral non-colour words that were matched for grapheme-length with the colour words (e.g., **DOG** in red ink or **JUMP** in blue) and presented in red, blue, green, or yellow. Stimuli in the congruent condition were colour-words that were presented in the colour congruent with their meaning (e.g., **RED** in red ink). Stimuli in the

incongruent condition were colour-words presented in any of the three colours that did not match their meaning (e.g., RED in blue ink).

Stimuli in the incongruent condition were further divided into semantically incongruent and response incongruent. Semantically incongruent stimuli contained semantic conflict only, as the response key was the same for both the target colour and the colour word (e.g., RED in blue ink; target colour and word are mapped onto the same left button). Response incongruent stimuli contained both semantic conflict and response conflict, as the response key for the target colour was different from the colour word (e.g., YELLOW in red ink; target colour is mapped onto the left button whereas word is mapped onto the right button). The response incongruent condition is the standard incongruent condition used in the traditional manual version of the Stroop task.

Procedure

A blocked experimental design was chosen for the study. Each experimental block lasted for 45 seconds and represented a single condition. A fixation block preceded the experimental block for 15 seconds. Stimuli were presented using E-prime (Psychological Software Tools, Pittsburgh, PA). There were 18 stimuli in each experimental block presented on a grey background. Each stimulus was presented for 2000 ms followed by 500 ms inter-stimulus interval, which was a blank grey screen. The stimulus was presented for 2000 ms across all trials regardless of whether a response was made during that time or not.

Participants were required to name the colour the word was presented in using a mouse as the response apparatus. Red and blue target colours were mapped onto the left mouse button whereas green and yellow were mapped onto the right mouse button. The participants were instructed to make a response while the stimulus was on the screen, and if they failed to make a response during that period, to wait for the next stimulus and respond to that. Reaction times and accuracy were recorded.

Data were collected over two runs. In each run a total of eight experimental blocks were included, with each condition being repeated. The order of conditions in each run was predetermined. The conditions in the first order were: control, congruent, semantically incongruent, response incongruent, response incongruent, semantically incongruent, congruent and control. The conditions in the second order were: response incongruent, semantically incongruent, congruent, control, congruent, semantically incongruent, and response incongruent. Each order was counterbalanced between the subjects. This method was employed to offset the effects of linear drifts in signal intensity which is inherent in fMRI acquisition.

Bilingual subjects completed a total of 4 runs, with language order being counterbalanced. All participants performed a practice version of the task, which included a single block of stimuli for each condition, to ensure good performance in the scanner.

Image acquisition

Images were acquired using a 1.5T Siemens Avanto scanner (Erlangen, Germany). Each scanning session started with an acquisition of a T1-weighted structural volumes using 3D MP-RAGE sequence (TR = 11ms; TE = 4.94 ms; flip angle: 15°; FOV: 25.6 x 20.8 cm; matrix size: 256x208; 170 to 176 axial slices parallel to AC-PC line; slice thickness: 1 mm; interslice gap: 0 mm; resulting in 1x1x1 mm voxels, axial acquisition, parallel to AC-PC line, ensuring whole brain coverage). At the beginning of each functional session three “dummy” scans were also acquired, which were not recorded, to control for signal saturation. A total of 162 volumes were acquired during each of the 2 functional scanning sessions, resulting in 324 T2*-weighted volumes per subject per language. The EPI acquisition sequence parameters were as follows: TR = 3000 ms; TE = 50 ms; flip angle = 90°; FOV= 19.2 cm; matrix size: 64x64; with interleaved slice acquisition, starting at the bottom; 30 slices parallel to AC-PC line; slice thickness: 4 mm; 25% gap: resulting in 3x3x5 mm voxels; whole brain coverage of 150 mm.

Image pre-processing and analysis

SPM5 software (Wellcome Department of Imaging Neuroscience, London, UK) was used for image processing and analysis. The first two image volumes were discarded, and the first volume of the first session was used as a reference for realigning the rest of the volumes and a mean of all volumes for the sessions was created. The T1-weighted structural image was coregistered to the mean of the functional volumes. Normalisation parameters were estimated using the unified segmentation procedure (Ashburner & Friston, 2005), and used to normalise the structural and functional images to the stereotactic coordinate system defined by the Montreal Neurological Institute (MNI). Finally, the functional volumes were spatially smoothed using an anisotropic Gaussian filter of 9 15 mm (3 times the voxel size) at full-width at half-maximum (FWHM).

Functional data were analysed in a two-step procedure, a within-subjects GLM time-series regression, and random-effects factorial ANOVA with the resulting parameter estimates as the dependent measure. The time-series analysis was performed using a general linear approach with the four experimental conditions (control, congruent, semantically incongruent and response incongruent). Each experimental block was modelled as a series of events, and each event was convolved with the canonical hemodynamic response function. The implicit baseline condition, the fixation screen, was not included in the model. Beta-estimates from GLM time-series analysis for the four experimental conditions were collapsed across the two runs. For each monolingual participant 4 contrast images were obtained whereas for each bilingual participant 8 contrast images

were obtained (4 for each language). These contrast images were imported into one of three factorial-design ANOVAs for further analysis.

Two 4 by 2 factorial-design ANOVAs with condition (control, congruent, semantically incongruent, response incongruent) as a within-subjects factor and group (monolinguals, bilinguals) as a between-subjects factor, one for English and one for L1, were evaluated. A 4 by 2 factorial-design ANOVA with condition and language (L1, L2) as within-subjects factors was evaluated for the bilingual group only. Correction for the violation of sphericity resulting from the use of repeated measures and possible heterogeneity of variance was implemented by SPM5. Model estimation was initially performed with all the available voxels which numbered 34 143 for the first two factorial ANOVAs and 34,859 for the third factorial ANOVA.

For each of the factorial ANOVAs only planned comparisons were performed, using an F-test. The first comparison looked at the response incongruent condition versus the congruent condition as this is a commonly used contrast in the literature for investigating the Stroop effect. The existing paradigm allowed for investigating an additional contrast that looked at the effect of response conflict only (response incongruent versus semantically incongruent). Uncorrected threshold of $p < .01$, and a contiguity threshold of 5 voxels was used for each comparison.

Results

Three separate analyses were performed on the accuracy and median reaction time data (RT; in milliseconds), and fMRI data. Firstly, the groups were compared in their performance in English. Secondly, the groups were compared

in their performance in their respective L1 and, in the third analysis, the bilinguals' performance was compared in their L1 (Macedonian) versus their L2 (English).

Behavioural results - English and L1

Both the RT and accuracy data were subjected to a Condition (control, congruent, semantically incongruent, response incongruent) \times Group (monolinguals, bilinguals) ANOVA, with Condition as a within-subjects factor.

The RT data analysis revealed no significant main or interaction effects, although the main effect of condition approached significance for English [$F_{(3,42)} = 2.447$, $p = .09$]. As expected, participants were slowest to respond to incongruent stimuli and fastest to congruent stimuli. Responses to control and semantically incongruent stimuli were very similar.

Overall, participants were fairly accurate across all conditions (above 90%). Mean and SD accuracies and RTs for each group across condition are shown in Table 1. The accuracy data analysis revealed no significant main or interaction effects, although the main effect of condition for L1 approached significance [$F_{(3,42)} = 2.758$, $p = .09$]. Participants were most accurate in the semantically incongruent condition, followed by the control, response incongruent and congruent condition respectively.

L1 vs. L2 for bilinguals

Both the RT and accuracy data were subjected to a Condition (control, congruent, semantically incongruent, response incongruent) \times Language (L1, L2) ANOVA, with Condition and Language as within-subjects factors. The RT data analysis revealed no significant main or interaction effects. Participants were fastest in

Table 1. Mean reaction times (RTs) in milliseconds and mean accuracy (ACC) for each condition (congruent, control, semantically incongruent, response incongruent) for monolinguals and bilinguals in their first language (L1) and second language (L2). Standard deviations are shown in parenthesis.

Condition	Congruent		Control		Semantically Incongruent		Response Incongruent	
	RT	ACC	RT	ACC	RT	ACC	RT	ACC
Monolinguals	514.1 (75.2)	.98 (.01)	498.7 (79.3)	.99 (.03)	504.6 (76.2)	.99 (.02)	536.6 (96.4)	.98 (.03)
Bilingual L1	573.7 (117.1)	.93 (.07)	634.9 (208.1)	.94 (.08)	573.1 (157.5)	.96 (.08)	604.8 (122.2)	.93 (.07)
Bilingual L2	543.9 (88.8)	.93 (.09)	573.4 (117.5)	.93 (.07)	560.1 (117.5)	.94 (.06)	579.4 (79.0)	.92 (.10)

Table 2. Brain regions showing significant group differences in English for each contrast, including cluster extent in mm³, Talairach coordinates for the peak activation voxel, summary statistics, Brodmann Area (BA), and direction of difference. More than one local maxima more than 8mm apart are shown and a single voxel extent is 45mm³.

Region	F	Cluster Extent mm ³	X Y Z	BA
Response incongruent vs. congruent: Monolinguals > Bilinguals				
(1) R. Superior Frontal G.	13.38***	450	18 11 50	6
(2) R. Middle Frontal G.	9.59**	225	45 3 51	6
R. Inferior Frontal G.	9.08**	405	50 23 -5	47
L. Fusiform G.	8.92**	270	-99	37
(3) R. Inferior Frontal G.	8.87**	225	50 15 18	44
L. Fusiform G.	8.83**	585	-103	19
(4) L. Cingulate G.	8.57**	495	-9 11 41	32
L. Cingulate G.			0 11 36	32
L. Lingual G.	7.9**	225	-24 -64 3	19
Bilinguals > Monolinguals				
(5) L. Caudate Tail	9.39**	315	-15 -31 20	
(6) L. Precentral G.	9.18**	495	-36 -15 42	4
Response incongruent vs. semantically incongruent: Monolinguals > Bilinguals				
(1) R. Insula	16.76***	225	33 4 18	13
(2) R. Anterior Cingulate	11.72***	405	18 38 -2	32
(3) R. Superior Frontal G.	9.3**	450	18 14 50	6
(4) R. Inferior Frontal G.	9.22**	225	50 15 18	44
Bilinguals > Monolinguals				
L. Pons	13.53***	225	-50	
L. Sub-Gyral	10.92**	270	-18 -6 56	6
(5) L. Parahippocampal G.	8.8**	270	-58	
L. Thalamus	8.59**	225	-15 -31 15	

the congruent condition, followed by the semantically incongruent, response incongruent and control conditions. The accuracy data analysis revealed no significant main or interaction effects. Overall, participants were fairly accurate across all conditions (above 90%).

Both the RT data (in the range of 500 - 600ms) and accuracy data (over 91%) indicate that the participants may have found the task relatively easy. All participants underwent an extensive practice session prior to going into the scanner and they all responded with their right hand on the mouse, most likely a

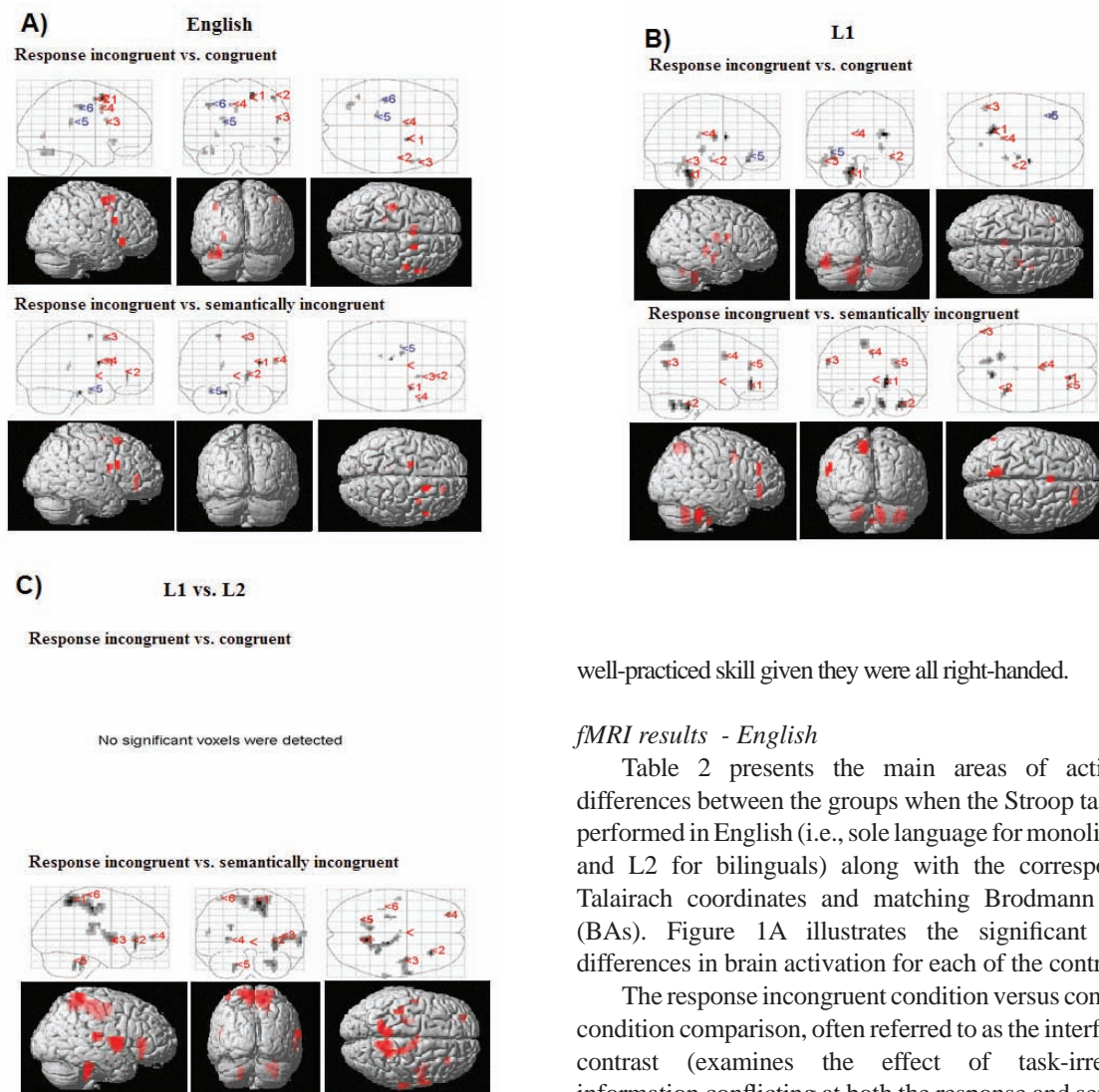


Figure 1. (A). Significant differences between groups in English for each contrast displayed on a cortical surface rendering and glass brain SPMs. Selected peak voxels within the clusters shown on the SPMs. Labels for shown voxels are presented in Table 2. Red numbers indicate regions of greater activation for monolinguals, whereas blue indicate greater activation for bilinguals. (B). Significant differences between groups in L1 for each contrast. Labels for shown voxels are presented in Table 3. Red numbers indicate regions of greater activation for monolinguals, whereas blue indicate greater activation for bilinguals; (C) Significant differences between two languages in bilinguals for each contrast. Labels for shown voxels are presented in Table 4. Red numbers indicate regions of greater activation for L2, whereas blue indicate greater activation for L1.

well-practiced skill given they were all right-handed.

fMRI results - English

Table 2 presents the main areas of activation differences between the groups when the Stroop task was performed in English (i.e., sole language for monolinguals and L2 for bilinguals) along with the corresponding Talairach coordinates and matching Brodmann Areas (BAs). Figure 1A illustrates the significant group differences in brain activation for each of the contrasts.

The response incongruent condition versus congruent condition comparison, often referred to as the interference contrast (examines the effect of task-irrelevant information conflicting at both the response and semantic levels), differed between the groups, with monolinguals showing greater activation in the right superior, middle, and inferior frontal gyri, and the left fusiform, lingual, and cingulate gyri. Bilinguals showed greater activation than monolinguals in the left caudate tail only.

The response incongruent condition versus semantically incongruent condition contrast (examines the effect of competing task-relevant information that conflicts at the response level) also differed between groups, with monolinguals showing greater activation in the anterior cingulate, superior and inferior frontal gyri, and insula in the right hemisphere. Bilinguals showed greater activation in the left parahippocampal gyrus, thalamus, and pons.

It is noteworthy that activation differences were observed between the groups in regions such as the ACC and DLPFC,

Table 3. Brain regions showing significant group differences in L1 for each contrast, including direction, cluster extent mm³, Talairach coordinates, summary statistics, Brodmann area (BA), and direction of difference. More than one local maxima more than 8mm apart are shown where appropriate. A single voxel extent is 45mm³.

Region	F	Cluster Extent mm ³	X Y Z	BA
Response incongruent vs. congruent: Monolinguals > Bilinguals				
(1) L. Cerebellum	18.17***	2745	-81	
R. Claustrum	17.81***	270	-33 28 -14	
R. Claustrum	12.56***	540	27 -8 19	
R. Thalamus	10.98**	450	27 -26 1	
(2) R. Temporal Lobe	9.84**	225	39 -15 -12	21
(3) L. Fusiform G.	9.35**	945	-112	37
(4) L. Posterior Cingulate	9.05**	225	0 -31 20	23
L. Cingulate G.			-3 -25 24	23
R. Cerebellum	8.51**	225	9 -56 -22	
Bilinguals > Monolinguals				
(5) L. Inferior Frontal G.	12.77***	630	-33 28 -14	47
Response incongruent vs. semantically incongruent: Monolinguals > Bilinguals				
(1) R. Anterior Cingulate	14.54***	810	18 38 -6	32
(2) R. Cerebellum	12.96***	810	36 -39 -27	
R. Cerebellum	12.73***	1170	12 -57 -27	
L. Cerebellum	11.15***	990	-90	
(3) L. Superior Temporal	10.47**	360	-56 -63 26	39
L. Precuneus	10.44**	1755	-9 -56 53	7
(4) L. Cingulate G.	10.05**	540	0 8 36	24
(5) R. Middle Frontal G.	9.63**	540	33 36 17	10
Bilinguals > Monolinguals				
R. Brainstem Medulla	7.8**	225	3 -28 -36	

Table 4. Brain regions showing significant differences between the languages for each contrast in bilinguals, including cluster extent in mm³, Talairach coordinates for the peak activation voxel, summary statistics, Brodmann area (BA), and direction of difference. More than one local maxima more than 8mm apart are shown. Note that a single voxel extent is 45mm³.

Region	F	Cluster Extent mm ³	X Y Z	BA
Response incongruent vs. semantically incongruent: L2 > L1				
(1) R. Precuneus	14.66***	6975	12 -47 58	7
R. Paracentral L.			15 -21 47	6
R. Superior Parietal L.			21 -41 62	5
(2) R. Inferior Frontal G.	12.4***	990	30 29 -1	47
R. Inferior Frontal G.			30 29 -10	47
(3) R. Insula	11.38***	3105	39 0 0	13
R. Superior Temporal G.			62 3 0	22
R. Superior Temporal G.			50 -3 5	22
L. Paracentral L.	11.25***	2385	-6 -41 57	5
L. Postcentral G.			-18 -49 67	7
R. Cerebellum	10.94**	1215	30 -39 -27	
R. Parahippocampal G.			21 -36 -23	36
R. Cerebellum			30 -40 -36	
(4) L. Middle Frontal G.	10.54**	360	-30 50 2	10
(5) L. Cerebellum	10.39**	360	-90	
L. Claustrum	9.96**	495	-24 -22 20	
R. Postcentral G.	9.32**	495	65 -23 15	40
(6) L. Precentral G.	8.87**	360	-36 -26 66	4
L. Precentral G.	8.58**	270	-50 -8 14	43

suggesting a greater role for these regions when dealing with conflicting information in monolinguals compared to bilinguals.

L1

Table 3 presents the main areas of activation differences between the groups during task performance in L1 (i.e., English for monolinguals and Macedonian

for bilinguals) along with the corresponding Talairach coordinates and matching BAs for each of the contrasts.

As shown in Figure 1B, differences in brain activation between the groups in L1 can be observed in a number of regions for each contrast. The response incongruent condition versus congruent condition analysis showed that monolinguals had greater activation

in the fusiform and cingulate gyri, and cerebellum in the left hemisphere, and the right cerebellum and temporal lobe. Bilinguals showed greater activation in the left inferior frontal gyrus only.

The response in congruent condition versus semantically incongruent condition analysis also resulted in group differences, with monolinguals showing greater activation in the right anterior cingulate, left superior temporal gyrus, left precuneus, left cingulate gyrus, left cerebellum, and in the right middle frontal gyrus and cerebellum. Bilinguals showed greater activation in the right brainstem.

L1 vs. L2

Presented in Table 4 are the main areas of activation differences between the languages for the bilinguals, for each of the contrasts, along with Talairach coordinates and matching BAs.

Figure 1C shows the results of the functional MRI analysis for the two contrasts. No differences were observed between the two languages for the response incongruent condition versus the congruent condition contrast.

The response incongruent condition versus semantically incongruent condition contrast differed between the languages, with L2 showing significantly greater activation than L1 in numerous areas in the right hemisphere (precuneus; insula; superior parietal lobule; inferior frontal gyrus; superior temporal gyrus; parahippocampal gyrus; and postcentral gyrus), in the left hemisphere (middle frontal gyrus, precentral gyrus), and bilaterally in the paracentral lobule and cerebellum.

Discussion

The purpose of this study was to examine functional neuroanatomy during Stroop colour-word task performance in late proficient bilinguals compared to monolinguals. The behavioural data showed the expected Stroop interference effect. Both the RT data (in the range of 500-600ms) and accuracy data (over 90%) indicate that the participants found the task relatively easy.

We designed the experiment to allow for comparisons both between groups and between languages on two cognitive components of the Stroop task: interference and response conflict. Although the interference contrast (i.e., the response incongruent versus congruent condition comparison) is the most common approach of evaluating the Stroop effect, we were particularly interested in the conflict resolution process. Conflict detection is thought to be carried out by the ACC (Botvinick et al., 1999), while the allocation of attentional resources required to resolve the conflict is thought to depend on prefrontal cortex (Bunge, Dudukovic, Thomason, Vaidya, & Gabrieli, 2002; MacDonald et al., 2000).

Group comparisons between monolinguals (in English) and bilinguals (in their first language of Macedonian)

revealed greater activation for monolinguals in the left cingulate (BA 32) and fusiform gyrus (BA 37). Greater activation by monolinguals in the posterior regions might reflect greater attention to, and more elaborate evaluation of, all stimulus dimensions of the Stroop stimuli (e.g., colour and meaning of the stimulus), given the association of these regions to stimulus processing, colour attention tasks, and reading and recognizing word forms (Cabeza & Nyberg, 2000; Clark et al., 1997; Dehaene, Le Clec, Poline, Le Bihan, & Cohen, 2002). It is also possible that less posterior activation observed in bilinguals might be indicative of less effortful word-form processing. A similar speculation was advanced by Kovelman (2006) in a semantic judgment task study with early proficient bilinguals.

Bilinguals showed greater activation than monolinguals in the left inferior frontal gyrus (BA 47) which is consistent with previous reports with early bilinguals during conflict tasks (Bialystok et al., 2005). The left inferior frontal gyrus is known to be involved in language production and comprehension, but it also plays a role in manipulating and associating information held in working memory (Petrides, 1995). Given this, a greater reliance on inferior frontal regions might be the general result of the extensive practice that bilinguals get manipulating two languages.

The comparison between monolinguals and bilinguals in English (bilingual L2) for the response conflict contrast revealed greater activation for monolinguals relative to bilinguals in the ACC, right inferior and superior frontal gyrus, and insula. This is consistent with earlier observations on this task with monolinguals by van Veen and Carter (2005). Monolinguals also showed more activation in the ACC and right middle frontal gyrus than bilinguals when comparisons included the bilinguals' native language. As the activation differences did not include areas involved in motor output, the group differences most likely relate to conflict resolution prior to the behavioural response.

We noted earlier that the ACC is involved in conflict detection, but it is also more generally thought to be associated with the monitoring of behaviour, maintenance of available responses in working memory, and inhibitory control of inappropriate responses (Carter et al., 1998; MacDonald et al., 2000; van Veen & Carter, 2002). The ACC may even be related to task difficulty (Barch, Carter, Hachten, Usher, & Cohen, 1997) and/or practice (Bench et al., 1993). As poor performance on the Stroop task has been associated with greater activation in anterior cingulate and prefrontal regions (Bush et al., 1999; Weiss et al., 2003) and older adults tend to recruit additional, particularly prefrontal, areas to maintain good Stroop task performance (Langenecker, Nielson, & Rao, 2004), our results might also indicate that bilinguals are more effectively dealing with response conflict than monolinguals. This is consistent with the idea that positive changes in executive functioning can

occur in bilingual individuals as a result of the constant management and use of the two languages (Bialystok, 2001; Costa et al., 2008).

The possibility that the reduced activation observed in bilinguals is related to more effective task performance at the neural level is a tentative possibility only, as reduced activation could also indicate less efficient cognitive control. Research with schizophrenia patients, for example, has found reduced activation in the anterior cingulate region compared to healthy controls (Kerns et al., 2005; Weiss et al., 2007).

Interestingly, our bilinguals showed significant activation during English task performance in the parahippocampal region. Activity in this region has been previously observed in late and proficient bilinguals during story listening in both languages (Perani et al., 1998), suggesting an increased participation of declarative-memory-related structures which are typically considered important for lexical-semantic knowledge (Paradis, 2004). However, the use of this system during Stroop task performance is less clear and requires further study.

The comparison between the two languages for bilinguals revealed greater activation in the left middle frontal gyrus and the right inferior frontal gyrus during performance in English (L2) than during Macedonian (L1). The greater involvement of prefrontal regions for L2 suggests that greater resources were required for second language processing and for monitoring task performance in English. Neuroimaging evidence along with lesion studies suggests that frontal regions are engaged during language switching and selection (e.g., Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001; Meuter, Humphreys, & Rumiati, 2002), and this experience may allow for refinement and application of these skills to other domains.

The insula was also significantly more activated in L2 than in L1. The insula, thought to play a key role in language processing and in the production of relatively “automatic” responses (Raichle et al., 1994), may act as a marker for language attainment in bilinguals (Chee, Soon, Lee, & Pallier, 2004). In the fMRI study by Chee et al. (2004), early and proficient bilinguals showed greater activation in the left insula and reduced activation in the anterior cingulate compared to less proficient bilinguals during a working memory task in an unfamiliar language.

Thus, in the current study, where bilinguals are living and working in their L2 environment, L2 might become more automatic than L1 over time. Indeed, environmental exposure to language has been shown to be an important factor in the cerebral representation and processing of language in bilinguals (Badzakova-Trajkov, Kirk, & Waldie, 2008; Perani et al., 2003). Over time, with language processing becoming less automatic in the native language, neural changes to structures

involved in both L2 and L1 might even occur.

One limitation of the current study is the lack of a monolingual Macedonian comparison group. This would have been useful to help rule out language-specific effects. As the study was conducted in an English speaking country, however, the inclusion of such group was not currently possible. A second limitation of this study is that using a two button response box might have made the task easier than anticipated. A more difficult task paradigm may be more appropriate for revealing group differences on tasks that require a high level of executive control. Third, a relatively relaxed alpha level was used for the comparisons of interest. A larger sample size would have been advantageous in terms of using a more stringent level.

Overall, the results showed that monolinguals produced greater activation than bilinguals during the Stroop task, regardless of the attentional task demands. More specifically, monolinguals showed more extensive posterior brain activation during Stroop interference, which might reflect more elaborate stimulus evaluation. Monolinguals also showed greater activation in the anterior cingulate and prefrontal regions during response conflict. For bilinguals, performance in English activated a greater number of neural substrates in general, suggesting that the second language demanded greater overall processing resources than L1. However, greater activation in areas such as the insula suggests that L2 is becoming more automatic and efficient, which might be expected considering that bilingual participants, despite learning English after entering primary school, have constant environmental exposure to their second language. Taken together, this study provides reasonable evidence that the neural substrates underlying Stroop task performance are different depending on whether or not a second language was acquired.

Acknowledgements

This research was supported by the University of Auckland Emerging Research Excellence Award (3602182) to the corresponding author. We wish to thank all our participants for their involvement in this study.

References

- Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: a meta-analytic review. *Neuropsychology Review*, *16*, 17-42.
- Ashburner, J., & Friston, K.J. (2005). Unified segmentation. *NeuroImage*, *26*, 839-851.
- Badzakova-Trajkov, G., Kirk, I.J., & Waldie, K.E. (2008). Dual-task performance in late proficient bilinguals. *Laterality*, *13*, 201-216.
- Barch, D.M., Carter, C.S., Hachten, P.C., Usher, M. & Cohen, J.D. (1999). The “benefits” of distractibility: Mechanics underlying increased Stroop effects in schizophrenia. *Schizophrenia Bulletin*, *25*, 749-762.

- Bench, C.J., Frith, C.D., Grasby, P.M., Friston, K.J., Paulesu, E., Frackowiak, R.S. et al. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, *31*, 907-922.
- Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*, 3-11.
- Bialystok, E., Craik, F.I.M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*, 290-303.
- Bialystok, E., Craik, F.I.M., Grady, C., Chau, W., Ishii, R., Gunji, A. et al. (2005). Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage*, *24*, 40-49.
- Bialystok, E., Craik, F.I.M., & Ryan, J. (2006). Executive control in a modified antisaccade task: Effects of aging and bilingualism. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1341-1354.
- Bialystok, E., & Shapero, D. (2005). Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science*, *8*, 595-604.
- Botvinick, M.M., Nystrom, L.E., Fissell, K., Carter, C.S., & Cohen, J.D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*, 179-181.
- Bunge, S.A., Dudukovic, N.M., Thomason, M.E., Vaidya, C.J., & Gabrieli, J.D.E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, *33*, 301-311.
- Bush, G., Frazier, J.A., Rauch, S.L., Seidman, L.J., Whalen, P.J., Jenike, M.A. et al. (1999). Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the counting Stroop. *Biological Psychiatry*, *45*, 1542-1552.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*, 1-47.
- Carlson, S.M., & Meltzoff, A.N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, *11*, 282-298.
- Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., & Cohen, J.D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*, 747-749.
- Clark, V.P., Parasuraman, R., Keil, K., Kulansky, R., Fannon, S., Maisog, J.M. et al. (1997). Selective attention to face identity and colour studied with fMRI. *Human Brain Mapping*, *5*, 293-297.
- Cohen, J.D., Botvinick, M., & Carter, C.S. (2000). Anterior cingulate and prefrontal cortex: Who's in control? *Nature Neuroscience*, *3*, 421-423.
- Costa, A., Albareda, B., & Santesteban, M. (2008). Assessing the presence of lexical competition across languages: Evidence from the Stroop task. *Bilingualism: Language and Cognition*, *11*, 121-131.
- Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology*, *26*, 1283-1296.
- Dehaene, S., Le Clec, H.G., Poline, J.B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *NeuroReport*, *13*, 321-325.
- Green, D.W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 67-81.
- Hernandez, A.E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish-English bilinguals: An fMRI study. *NeuroImage*, *14*, 510-520.
- Kerns, J.G., Cohen, J.D., MacDonald III, A.W., Johnson, M.K., Stenger, V.A., Aizenstein, H. et al. (2005). Decreased conflict- and error-related negativity in the anterior cingulate cortex in subjects with schizophrenia. *American Journal of Psychiatry*, *162*, 1833-1839.
- Kovelman, I. (2006). *Bilingual and monolingual brains compared: A fMRI study of semantic processing*. Unpublished Doctoral Dissertation, Dartmouth College, Hanover, New Hampshire.
- Langenecker, S.A., Nielson, K.A., & Rao, S.M. (2004). fMRI of healthy older adults during Stroop interference. *NeuroImage*, *21*, 192-200.
- Lezak, M.D. (1995). *Neuropsychological Assessment*. New York: Oxford University Press.
- MacDonald III, A.W., Cohen, J.D., Stenger, V.A., & Carter, C.S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835-1838.
- MacLeod, C.M., & MacDonald, P.A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, *4*, 383-391.
- Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing: Converging evidence from eyetracking and brain imaging. *Brain and Language*, *86*, 70-82.
- Martin-Rhee, M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *11*, 81-93.
- Meuter, R.F.I., Humphreys, G.W., & Rumiati, R.I. (2002). The frontal lobes and bilingual language switching: Modulatory control in language selection. *International Journal of Bilingualism*, *6*, 109-124.
- Milham, M.P., & Banich, M.T. (2005). Anterior cingulate cortex: An fMRI analysis of conflict specificity and functional differentiation. *Human Brain Mapping*, *25*, 328-335.
- Miller, E.K., & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202.
- Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97-113.
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins Publishing Company.
- Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S.F. et al. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: An fMRI study during verbal fluency. *Human Brain Mapping*, *19*, 170-182.
- Perani, D., Paulesu, E., Galles, N.S., Dupoux, E., Dehaene, S., Bettinardi, V. et al. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain*, *121*, 1841-1852.
- Petrides, M. (1995). Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *Journal of Neuroscience*, *15*, 359-375.
- Petrides, M. (2000). Dissociable roles of mid-dorsolateral prefrontal and inferotemporal cortex in visual working memory. *The Journal of Neuroscience*, *20*, 7496-7503.
- Posner, M.I., & Dehaene, S. (1994). Attentional Networks. *Trends in Neurosciences*, *17*, 75-79.
- Raichle, M.E., Fiez, J.A., Videen, T.O., MacLeod, A.M., Pardo, J.V., Fox, P.T. et al. (1994). Practice related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, *4*, 8-26.
- Roelofs, A., Van Turenout, M., & Coles, M.G.H. (2006). Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *Proceedings of the National Academy of Sciences USA*, *103*, 13884-13889.
- Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.
- van Veen, V., & Carter, C.S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & Behaviour*, *77*, 447-482.
- van Veen, V., & Carter, C.S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *NeuroImage*, *27*, 497-504.
- Wager, T.D., Jonides, J., & Reading, S. (2004). Neuroimaging studies of shifting attention: A meta-analysis. *NeuroImage*, *22*, 1679-1693.
- Weiss, E.M., Golaszewski, S., Mottaghy, F.M., Hofer, A., Hausmann, A., Kemler, G. et al. (2003). Brain activation patterns during a selective attention test- a functional MRI study in healthy volunteers and patients with schizophrenia. *Psychiatry Research*, *123*, 1-15.
- Weiss, E.M., Siedentopf, C., Golaszewski, S., Mottaghy, F.M., Hofer, A., Kremser, C. et al. (2007). Brain activation patterns during a selective attention test- a functional MRI study in healthy volunteers and unmedicated patients during an acute episode of schizophrenia. *Psychiatry Research: Neuroimaging*, *154*, 31-40.