NeuroImage **13,** 1155–1163 (2001)
doi:10.1006/nimg.2001.0781, available online at http://www.idealibrary.com on IDEAL®

Relative Language Proficiency Modulates BOLD Signal Change when Bilinguals Perform Semantic Judgments

Michael W. L. Chee, Nicholas Hon, Hwee Ling Lee, and Chun Siong Soon

Cognitive Neuroscience Laboratory, Singapore General Hospital, 20 College Road, Singapore 169856, Singapore

Received November 6, 2000

The effect of relative language proficiency on the spatial distribution and magnitude of BOLD signal change was evaluated by studying two groups of right-handed English–Mandarin bilingual participants with contrasting language proficiencies as they made semantic judgments with words and characters. Greater language proficiency corresponded to shorter response times and greater accuracy in the semantic judgment task. Within the left prefrontal and parietal regions, the change in BOLD signal was smaller in a participant's more proficient language. The least proficient performance was associated with right, in addition to left, inferior frontal activation. The results highlight the importance of taking into consideration nature of task and relative language proficiency when drawing inferences from functional imaging studies of bilinguals.

Key Words: **Chinese characters; fMRI; semantic memory; bilinguals; language proficiency.**

INTRODUCTION

Functional imaging studies have shown that blood flow changes occur in spatially congruent regions when proficient or relatively proficient bilinguals perform linguistic tasks in different languages (Chee *et al.*, 1999a, 1999b, 2000; Illes *et al.*, 1999; Klein *et al.*, 1999; Perani *et al.*, 1998). In these studies, the spatial pattern of activation appears to be similar across languages irrespective of the surface features of the languages compared. As there is evidence that proficient bilinguals can access concepts directly from second language (L2) without having to perform an internal translation through first language (L1) (Kroll and deGroot, 1997), we suggested that common conceptual access is matched by the overlap of neuronal networks for processing L1 and L2 (Chee *et al.*, 1999a).

There is less agreement regarding less proficient bilinguals. Dehaene *et al.* (1997) reported inter- and intrahemispheric differences in spatial location when participants listened to sentences. Perani *et al.* (1996, 1998) reported a greater extent of activation for L1 when participants listened to sentences presented in L1 or L2, an effect that disappeared when participants were matched for proficiency in both languages. Finally, marginally more prefrontal activation in the participants' less proficient language has also been observed (Hernandez *et al.*, 2000).

We previously observed that Singaporeans (SGP) showed higher BOLD (blood oxygen level dependent) signal change when they performed a semantic judgment task in Mandarin, compared to when they did the same task in English (Chee *et al.*, 2000). This finding differed from a previous study that drew from the same population but which did not show cross-language differences in activation while the participants evaluated sentence meaning (Chee *et al.*, 1999a). In the study involving sentences, performance was closely matched across languages, whereas in the associative semantics experiment, performance was slower and less accurate in Mandarin, the participants' less proficient language. A plausible explanation for the contrasting observations is to attribute the differences in activation (and performance) to the relative proficiencies of L1 and L2. This effect of differential proficiency may have been evident in one task and not the other because of task-related-factors (e.g., overall difficulty of task).

An alternative explanation is that there is a real difference in the processing demands between these languages, whereby Mandarin inherently requires more resources to process, and that this difference was somehow not revealed in previous experiments.

Chinese (Mandarin and Chinese are used interchangeably here) remains the only major language to use a purely logographic script. It has been proposed that this characteristic makes the concept represented by each character (or group of characters) relatively transparent to the reader (Wang, 1973). While it is true that the semantic radical in Chinese provides a clue to the meaning of the character, the extent to which the semantic radical is related to the concept actually denoted by the complete character varies from vague to highly salient (Chen, 1999). Further, with common words in Modern Chinese, a good number of concepts

1053-8119/01 $35.00

are conveyed with two or more characters. In some of these compounds, the meaning of the compound is distinct from the meaning conveyed by its individual constituent characters. For example, "sheng qi," a compound meaning "angry" is derived from individual characters meaning "to give birth to" (or "to produce") and "air." A number of compounds have their origins in literary constructs that in times past had very salient semantic connections but are generally non-informative to the modern reader. It has also been suggested that the "phonetic" radicals in characters help the reader predict phonology. In reality, only about 25% of characters conform to such rules (DeFrancis, 1984). This figure is even lower for the most commonly used characters. Taken together, these factors could make logographs more difficult to read than alphabetic scripts. Finally, a recent imaging study contrasting covert verb generation in response to visually presented nouns by Mandarin–English bilinguals and native English speakers suggested that activation of part of the left dorsolateral prefrontal region (BA9) was present only in those with exposure to Chinese (Xiong *et al.*, 2000).

In order to discriminate between relative proficiency and inherent difficulty as modulators of cortical activation, we recruited two groups of English–Mandarin bilingual volunteers, one more proficient in English and the other more proficient in Mandarin. The expectation was that if the relative differences in BOLD signal magnitude were reversed in the group more familiar with Mandarin (i.e., more BOLD in English than in Mandarin), we could infer that language proficiency has a greater influence on BOLD response. If, on the other hand, greater activation for Mandarin was maintained across the two populations, we could infer that the processing demands of the two languages differ in a language-dependent manner.

METHODS

Singaporean (SGP) Participants

Ten right-handed, English–Mandarin bilingual participants (undergraduates or graduates) aged between 19 and 29 years (comprising six men and four women), gave informed consent for this study. These Singaporean students of Chinese ethnicity were exposed to both English and Mandarin at or before the age of 5 years. To gain entry into local undergraduate education, a passing grade in both languages is an important condition. Participants were chosen on the basis that they scored excellent grades in both languages in standardized high school examinations and that they use both languages in conversation in daily life. However, there is a bias towards English use, as this is the medium of instruction and commerce. Also, there is a very high level of exposure to English via the broadcast media, print media and Internet.

People's Republic of China (PRC) Participants

Nine right-handed, Mandarin–English bilingual participants (graduates who were either scientists or educators), five men and four women, aged between 23 and 34 years gave informed consent for this study. All the participants were exposed to Mandarin in elementary school while they were in China. They began learning to read and write English only in middle school, at or after the age of 12 years. These participants moved to Singapore only after completing college education in China, where Mandarin is the predominant medium of instruction and social interaction. In Singapore, they are exposed to English on a daily basis, either as the language of work or while pursuing their studies.

Behavioral Experiment

Details of the experimental tasks and presentation schema have been previously described (Chee *et al.*, 2000). The experiment consisted of a blocked design where stimulus triplets were presented for 3.0 s, followed by 0.5 s of fixation. Participants performed two different matching-to-sample tasks with English words or Mandarin characters. In the semantic task, they were instructed to choose the item closest in meaning to the sample stimulus. This task is known as the Pyramids and Palm Trees (PPT) task (Howard and Patterson, 1992) (Fig. 1). We used roughly similar numbers of natural and man-made items to ensure that any differential activation of category specific systems did not confound comparisons between characters and words (Moore and Price, 1999). In the size judgment tasks that served as the control conditions, one of the items was 6% smaller or larger than the sample item and the other was 12% smaller or larger. Participants were instructed to choose the item that was closest in size to the sample. Each of four experimental runs comprised two cycles of alternating word and character tasks and the order of presentation of word and character tasks was counterbalanced across runs (WCWC or CWCW). Response Time (RT) and accuracy data were collected while participants were scanned using a MR compatible, two-button mouse. After the scan session, each participant repeated the experiment in a relaxed and non time-constrained manner outside the scanner. Following this, they were instructed to indicate the words they could not identify. Analysis of the behavioral data was performed using a two-way repeated measures ANOVA (SPSS v10.0, SPSS Inc, Chicago, IL).

Imaging and Image Analysis

Scanning was performed in a 2.0T Bruker Tomikon S200 system (Bruker, Karlsruhe, Germany). A blipped gradient-echo EPI sequence with the following parameters was used: Effective TE 40 ms, TR 2000 ms, FOV 23 × 23 cm, a 128 × 64-pixel matrix. Fifteen oblique axial slices approximately parallel to the AC–PC line 4 mm thick (2-mm gap) were acquired. High-resolution anatomical reference images were obtained using a three-dimensional spoiled-gradient-recalled-echo sequence. Functional images underwent phase correction prior to further processing which was performed using Brain Voyager 2000 software (Brain Innovation, Maastricht, Holland). Intensity normalization was performed and followed by motion correction. Gaussian filtering was applied in the temporal and spatial domains. In the spatial domain a smoothing kernel of 4 mm FWHM was used. In the temporal domain, a 3 time-point FWHM filter was used. Registration of the functional MR data set to the high-resolution anatomical image of the brain was performed by manually registering the stack of T2 images acquired in an identical orientation to the functional MR data set to the 3-D image. The resulting realigned data set was then transformed into Talairach space (Talairach and Tournoux, 1988).

A linear cross correlation map of the size judgment tasks in each experiment was first performed. No significant differences emerged from this comparison.

Statistical maps were then computed using a general linear model (GLM) using two explanatory variables: character semantics and word semantics. The expected BOLD signal change was modeled using a gamma function (tau of 2.5 s and a delta of 1.5) synchronized to blocks of cognitive tasks. Statistical maps for individual participants were created using a correlation coefficient cut-off of 0.4 corresponding to an F value of approximately 100 (this corresponds to a P value of $<$ 0.001). Clusters of voxels smaller than 6 voxels were not displayed. This analysis resulted in the detection of activation in the prefrontal, left temporal, medial frontal, and left parietal regions. A multiple subject GLM was used to compute the pooled activation maps for each language and each cross language comparison in the SGP and PRC groups. Lenient correlation coefficient thresholds of 0.16 (SGP) and 0.18 (PRC) were used to accommodate the effect of between-subject variation in location of activation. (Note that the cc values are different on account of the different sample sizes and were chosen so that the calculated F values were approximately similar.) Talairach transformed group data was displayed on a volume-rendered brain of an individual from the cohort. Activations to a depth of 12 mm were projected to the surface and displayed using a red (min), yellow (max) color scale. (It is important to appreciate this as sulcal and opercular activation appear as if they are on the surface when, in reality they are not. However, this seems to be a widespread practice and does aid the visualization of activation.)

For each individual's data, regions of interest (ROI) in the left prefrontal and superior parietal regions were defined by sampling volumes that were active in both semantic comparisons (character or word) relative to size judgment. These areas were chosen for further evaluation of the relationship between behavioral measures and BOLD response because they yielded the most robust responses across participants. The cluster of voxels, bounded by a 3 × 3 × 3-cm cube centered around the activation peak defined the ROI (only activated voxels within this bounding cube were counted). Within each individual's ROIs, averaged time courses comprising 28 time points (14 task related and 14 baseline points) were calculated in order to show the average BOLD signal change due to the semantic tasks with respect to their size judgment baseline tasks. BOLD signal change was expressed as percentage signal change relative to the baseline (size judgment) task. Mean percentage signal change for each semantic judgment task was calculated from points 5 to 14 located on the plateau of the BOLD response corresponding to the semantic task and from the points in time 20 to 28 corresponding to the size judgment task (Fig. 4). In this way, points in the transition phase during the rise and fall of the BOLD signal were omitted. In order to determine if the differences in BOLD signal at particular ROIs could be generalized, the mean percent signal-change data from each subject was pooled for each population. The BOLD signal change elicited by the different semantic tasks in each language was then assessed using a two-tailed t test.

RESULTS

Behavioral Data

Self-generated reports of language usage and proficiency showed contrasting patterns. For PRC, Mandarin was the dominant language, whereas English was the dominant language for SGP.

Recognition data collected post-experiment indicated that the participants were familiar with most of the words and characters used in this study. On average, PRC reported that they could recognize 99.9% of the Mandarin characters and 90.1% of the English words presented. SGP reported that they could recognize 99.4% of the Mandarin characters and 99.5% of the English words.

For the untimed semantic task (Fig. 1), SGP were more accurate in English than in Mandarin (Fig. 2a), whereas the PRC were more accurate in Mandarin, as compared to English. A between-group comparison indicated that SGP performed better in their stronger

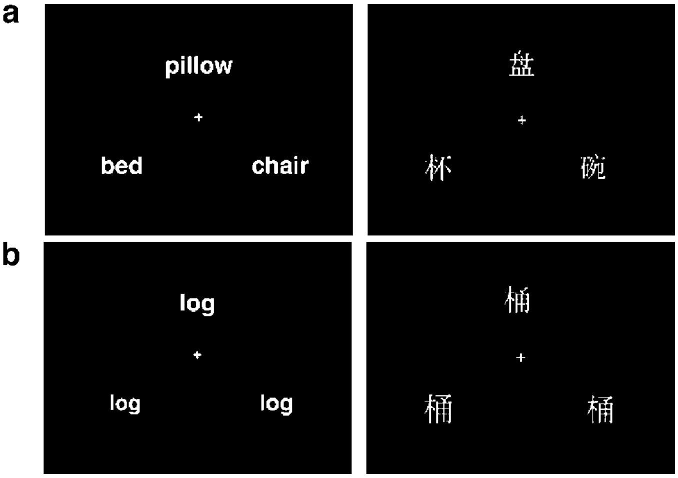

FIG. 1. Exemplars of the experimental task: (a) Semantic judgment and (b) size judgment in English and Mandarin.

language (i.e., English) than the PRC did in theirs (i.e., Mandarin). Both groups may be considered bilinguals, from the point of view of being able to identify concrete words and knowing enough about their properties to make associative judgments. However, the performance difference across the dominant languages suggests that, overall, the semantic network for PRC is slightly different than that of SGP, perhaps due to the influence of sociocultural factors that may influence the organization of the network. On the average, PRC appear to be at a slight disadvantage when performing semantic judgments on the test items used.

In the semantic judgment task performed under time constraint in the scanner, participants performed better when the task was presented in their dominant language: PRC were more accurate in the Mandarin version of the semantic judgment task, whereas SGP were better in English (Fig. 2b). Similarly, PRC were faster to respond when the task was presented in Mandarin, whereas SGP were faster when it was presented in English (Fig. 2c).

Imaging Data

As in previous work (Chee *et al.*, 2000), a network of left hemisphere predominant regions was revealed in the contrast between semantic and size judgment within individuals. This included the left prefrontal (BA 9, 44, 45), midline frontal (anterior SMA), left mid/posterior temporal (BA 21, 22), inferior temporal (BA 37) and left parietal (BA 7) regions (Table 1).

In SGP, both at the individual as well as group levels, there was no region active in one language that was spatially distinct from those activated in the other language despite the relative difference in proficiency. In the group data (z transformed to normalize individual contributions) involving all participants, subtraction of English semantic from Mandarin semantic ac-

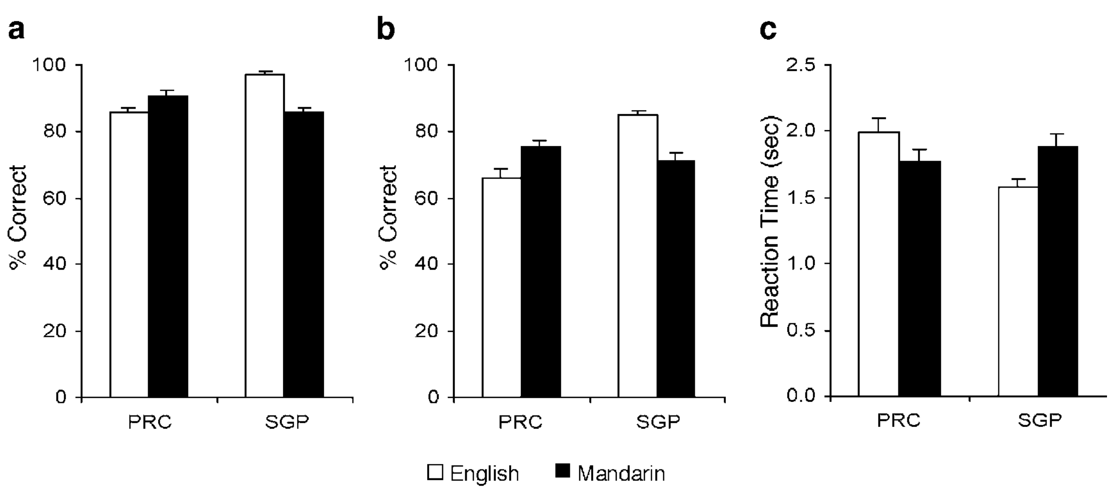

FIG. 2. (a) Accuracy for the semantic judgment task performed without time constraint. Overall, participants were more accurate in the English semantic judgment task [$F(1, 17) = 6.35$, $P < 0.05$], and SGP performed better than PRC [$F(1, 17) = 5.34$, $P < 0.05$]. Accuracy depended on the interaction between the task language and the nationality of the participants [$F(1, 17) = 40.09$, $P < 0.0001$]. Post-hoc t tests indicated that SGP performed better in the English version of the task [$t(9) = 6.01$, $P < 0.0001$], compared to the Mandarin version. PRC performed better in the Mandarin version of the task, $t(8) = -2.87$, $P < 0.05$. A between-group comparison indicated that SGP performed better in their stronger language (i.e., English) than the PRC did in theirs (i.e., Mandarin), $t(17) = -3.58$, $P < 0.005$. (b) Accuracy for the semantic judgment task performed with time constraint. There was no main effect of language, although SGP were more accurate than PRC [$F(1, 17) = 7.16$, $P < 0.05$]. Accuracy was dependent on the interaction between nationality and task language [$F(1, 17) = 38.60$, $P < 0.0001$]. Post-hoc within-group t tests confirmed that PRC were more accurate in the Mandarin version of the task [$t(8) = -2.95$, $P < 0.05$] and that SGP were more accurate in the English task [$t(9) = 6.79$, $P < 0.0001$]. (c) Response-time data for the semantic judgment task. There were no main effects of nationality or task language. Response time was dependent on the interaction of nationality and task language [$F(1, 17) = 90.57$, $P < 0.0001$]. Post-hoc within-group t tests indicated that PRC were faster to respond in the Mandarin version of the semantic judgment task [$t(8) = 6.86$, $P < 0.0001$]. SGP were faster in the English language task [$t(9) = -7.03$, $P < 0.0001$]. Error bars indicate 1 SEM.

TABLE 1

Talairach Coordinates of Activation Peaks in the Group Data

	SGP				PRC			
Brain region	*x*	*y*	*z*	*cc*	*x*	*y*	*z*	*cc*
English								
L Anterior cingulate/SMA	−5	29	44	0.18	−4	12	48	0.48
L Middle frontal gyrus (BA 9/44)	−36	7	30	0.38	−40	6	32	0.52
L Frontal operculum (BA 45)	−38	28	4	0.2	−41	19	6	0.4
R Frontal operculum (BA 45)	—	—	—	—	29	24	6	0.4
L Parietal (BA 7)	−24	−62	35	0.2	−24	−77	31	0.38
L Inferior temporal gyrus (BA 37)	−44	−54	−13	0.2	−49	−58	−9	0.22
Mandarin								
L anterior cingulate/SMA	−4	25	40	0.26	−3	16	46	0.4
L Middle frontal gyrus (BA 9/44)	−37	7	30	0.54	−39	12	31	0.44
L Frontal operculum (BA 45)	*	*	*	*	−29	28	15	0.34
L Parietal (BA 7)	−26	−61	40	0.28	−27	−74	38	0.36
L Inferior temporal gyrus (BA 37)	−42	−58	−17	0.28	—	—	—	—
L Middle temporal gyrus (BA 21/22)	−52	−45	−2	0.24	—	—	—	—
Mandarin > English								
L Anterior cingulate/SMA	−3	25	41	0.22	—	—	—	—
L Middle frontal gyrus (BA 9/44)	−37	6	30	0.42	—	—	—	—
L Frontal operculum (BA 45)	−29	24	4	0.22	—	—	—	—
L Parietal (BA 7)	−26	−64	45	0.24	—	—	—	—
L Inferior temporal gyrus (BA 37)	−41	−59	−17	0.24	—	—	—	—
English > Mandarin								
L Anterior cingulate/SMA	—	—	—	—	−3	8	48	0.38
L Middle frontal gyrus (BA 9/44)	—	—	—	—	−39	4	30	0.36
L Frontal operculum (BA 45)	—	—	—	—	−31	23	4	0.32
R Frontal operculum (BA 45)	—	—	—	—	30	24	6	0.38

Note. The correlation coefficient (cc) across groups are somewhat different because of the difference in sample size. The asterisk denotes that there was no distinct peak in the left ventral prefrontal region as activation was contiguous with that seen in the dorsolateral prefrontal cortex.

tivation revealed a pattern of left hemisphere areas that resembled the activation attributable to Mandarin but with the statistical maps generated using a higher threshold. No areas were more strongly activated for English compared to Mandarin (Fig. 3).

PRC displayed a broadly similar pattern of activation to that seen in SGP with several notable differences. In English, the less proficient language, additional areas of activation were present in the left and right opercular (including inferior frontal) regions. The additional activation was robust (i.e., survived elevation of the detection threshold) and was present in six of the nine participants. We interpret this finding as indicative of the recruitment of additional brain regions to perform semantic judgment in the less proficient language. Activation of the left posterior middle or superior temporal gyrus was observed in six of nine PRC but showed greater variability in spatial distribution than in SGP (8 of 10 demonstrated activation). As a result, temporal activation was not above threshold in the group data (Fig. 3) even though it was observed at the individual level (Fig. 4).

At the individual level, in both populations, and in the left prefrontal and parietal regions, the participants' more proficient language was associated with relatively smaller BOLD signal change than his or her less proficient language (Fig. 5). Across participants, the asymmetry index of BOLD activation during semantic judgment (taken across both languages) was highly correlated with the corresponding index of response times (Fig. 6).

DISCUSSION

The present results demonstrate that while processing the same linguistic task in different languages, BOLD signal is commensurate with proficiency. Specifically, a link between lesser proficiency, longer processing times and greater BOLD signal change in the left prefrontal and parietal areas is suggested. The results do not support the supposition that Mandarin is intrinsically more difficult to process for meaning than English.

Work to date suggests that within the limits of capable performance, RT and BOLD signal magnitude are linked to cognitive load as well as task specific processing requirements. RT, in tasks that tap common cognitive processes but vary in load, correlates directly with BOLD signal. Specifically, longer RT is associated with greater BOLD activity. In a pair of

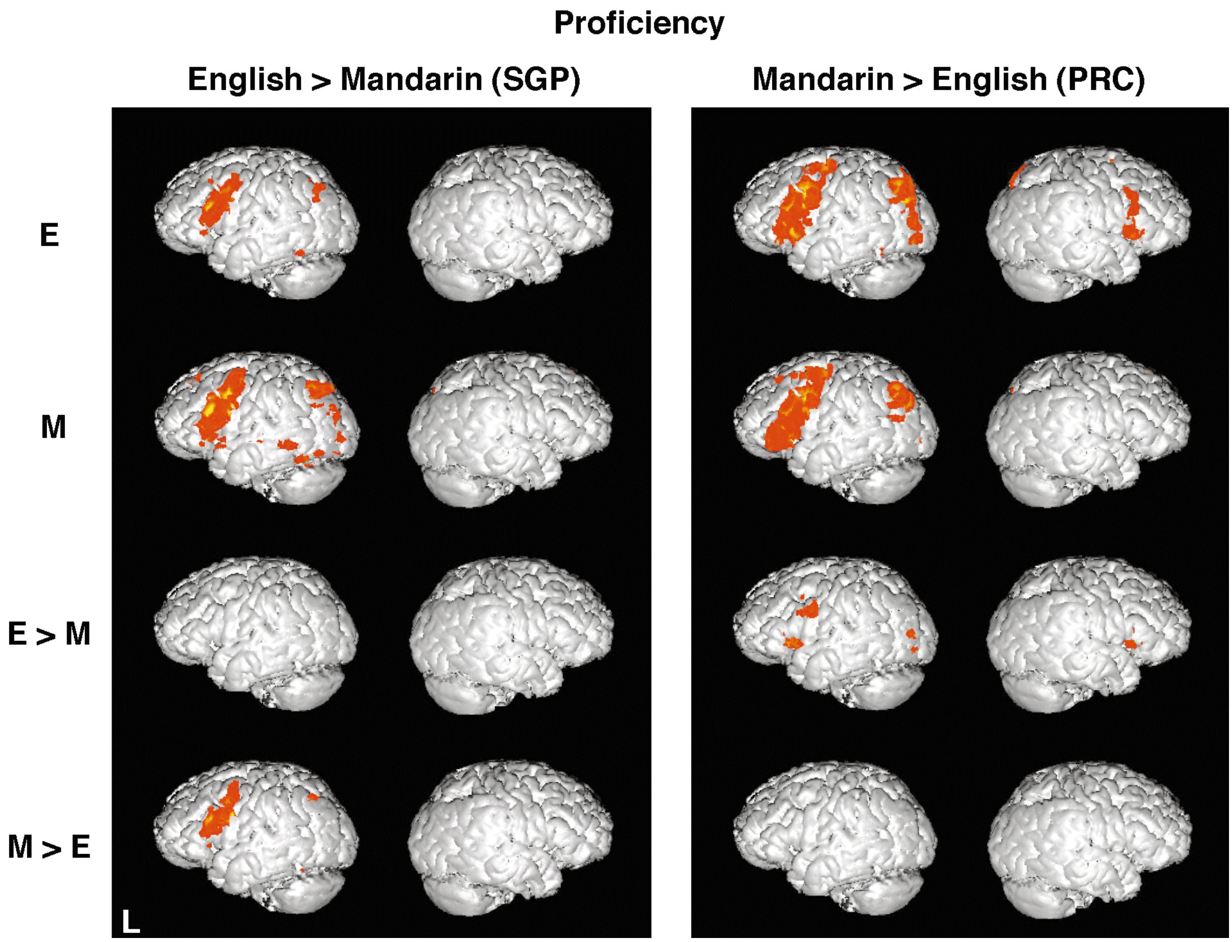

FIG. 3. Similarities and differences in activation associated with English (E) and Mandarin (M) semantic judgments in SGP and PRC participants. "Proficiency" refers to the relative language proficiencies of the two populations.

semantic comparison tasks, the high selection variant (that constrained which rule to use in classifying words) was associated with greater left prefrontal activation compared to the low selection variant (Thompson-Schill *et al.*, 1997). Prefrontal BOLD activation has also been shown to vary parametrically with increased working memory load (Rypma *et al.*, 1999; Smith and Jonides, 1997). In contrast, RT differences across tasks with different cognitive demands may be independent of BOLD signal differences (Demb *et al.*, 1995). Two perceptual judgment tasks involving consonant strings, one associated with a longer RT than a semantic word-classification judgment and another associated with a shorter RT were run. In the comparison against both "more difficult" and "less difficult" perceptual judgments, the semantic task was associated with a similar BOLD response in the left prefrontal region.

Given that the nature of our experimental task was the same in both languages, we propose that the increased RT in the participants' less proficient language is due to the greater cognitive effort needed to process words in the less familiar language.

Neurons in the prefrontal cortex respond differently to familiar compared to unfamiliar items. Specifically, these neurons fire in a more spatially restricted manner and less frequently in response to visual representations of familiar (relative to novel) pictures (Rainer and Miller, 2000). Although fewer neurons are involved in the recognition of familiar images, they are more finely tuned to identifying even degraded representations of these images. As such, one might suggest that familiar items require less "neuronal effort" for recognition. In the context of this study, words in the less familiar language may have less well-tuned representations, requiring greater neuronal activity (or "effort") for concept retrieval and manipulation. Since higher neuronal firing rates have been linked to greater BOLD signal (Rees *et al.*, 2000), lesser familiarity with a language may be related to greater BOLD activation.

There are at least two ways to interpret the additional bilateral opercular (and inferior frontal) activation in the PRC group. One could argue that these areas are recruited as a result of reduced plasticity in PRC consequent on their late acquisition of English.

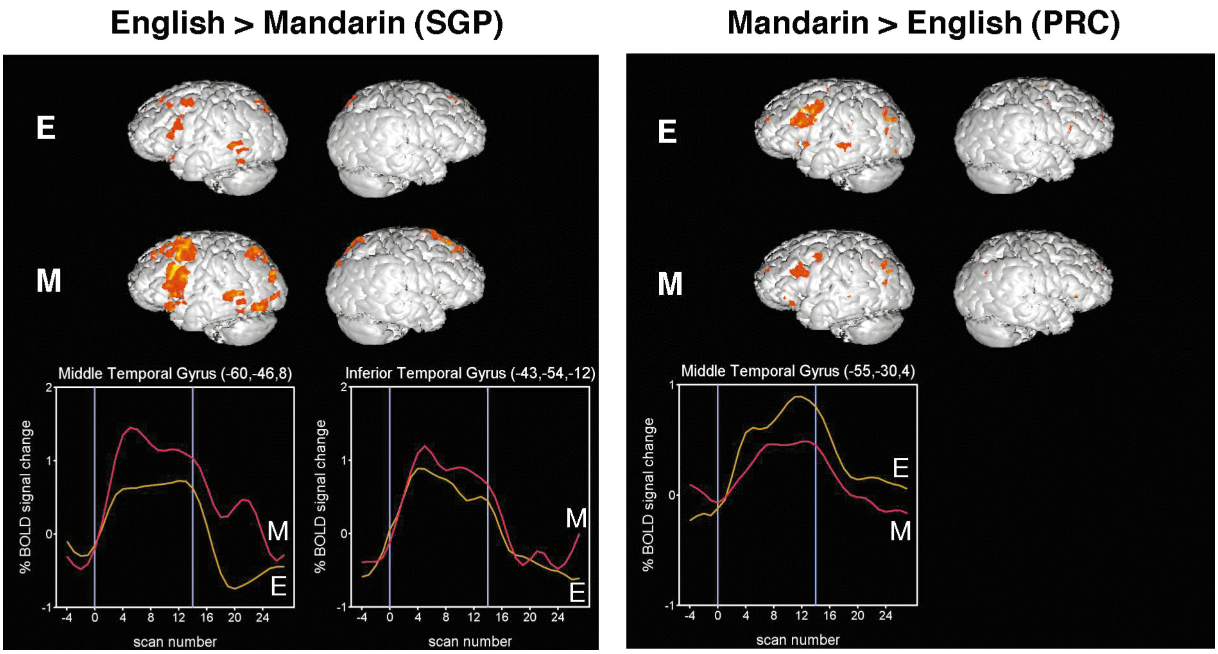

FIG. 4. Functional activation with the semantic task in English (E) and Mandarin (M) in two representative subjects drawn from SGP and PRC groups. The within-subject averaged time courses show percentage change BOLD signal in the left temporal region whose Talairach coordinates are displayed.

The findings of at least two imaging studies suggest that brain plasticity in second language acquisition may be limited. The peaks of left prefrontal activation were separated when "late bilinguals" silently thought about what they did at different times in the day in different languages whereas the prefrontal peaks overlapped in "early bilinguals" (Kim *et al.*, 1997). Also, recruitment of right hemisphere areas for American Sign Language (ASL) processing appears to be constrained by the earliness of acquisition (Bavelier *et al.*,

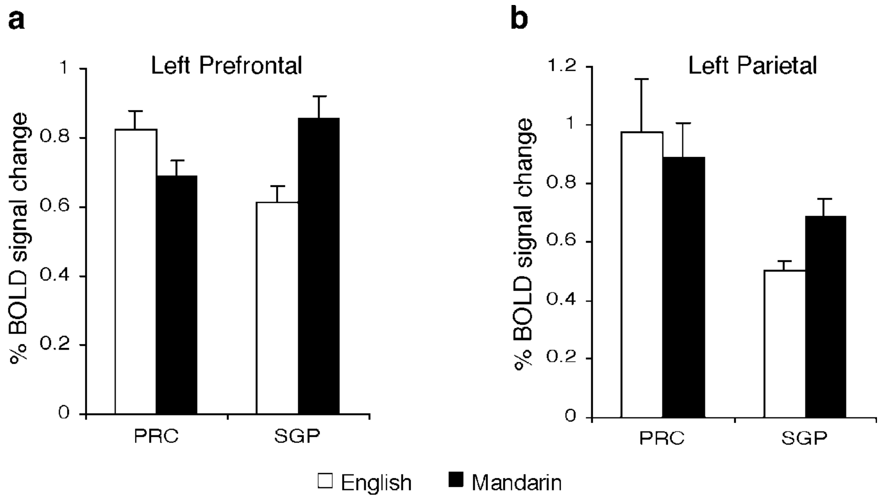

FIG. 5. (a) Left prefrontal BOLD activation for the semantic judgment task. There were no main effects of nationality or task language. BOLD activation was dependent on an interaction between task language and nationality [$F(1, 17) = 35.24$, $P < 0.0001$]. Post-hoc within-group t tests confirmed that SGP had less activation for the semantic judgment task performed in English, their stronger language [$t(9) = -4.72$, $P < 0.005$]. PRC had less activation for the task performed in Mandarin [$t(8) = 5.44$, $P < 0.01$]. (b) Parietal BOLD activation for the semantic judgment task. There was no main effect of language; however, PRC displayed greater activation in this region than their SGP counterparts [$F(1, 17] = 5.25$, $P < 0.05$]. BOLD activation was dependent on the interaction between task language and nationality [$F(1, 17) = 5.77$, $P < 0.05$]. Post-hoc within-group t tests revealed that there was no difference in BOLD when PRC performed the task in either language. However, there was more BOLD activity when SGP performed the task in Mandarin, compared to when they did it in English [$t(9) = -2.90$, $P < 0.05$]. Error bars indicate 1 SEM.

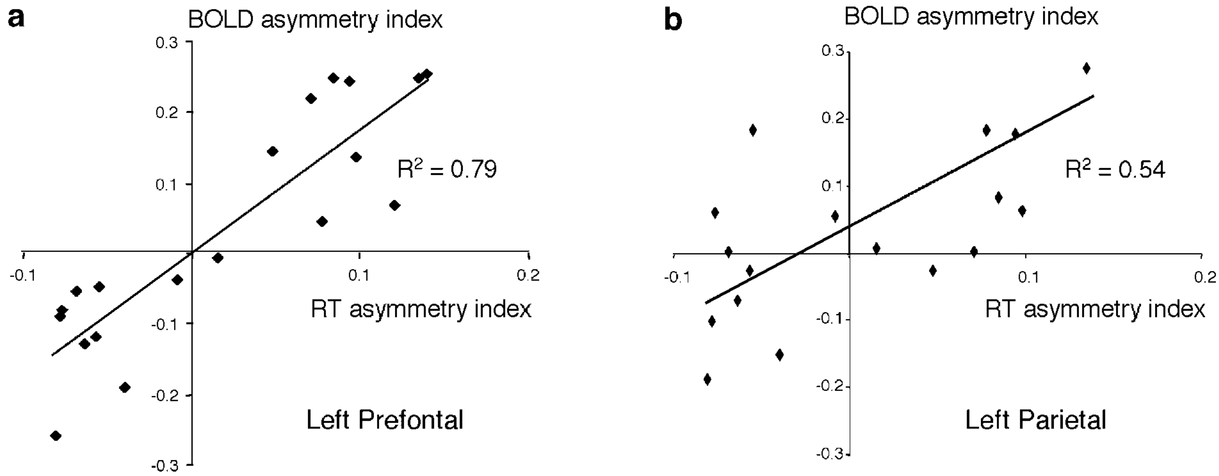

FIG. 6. Regression plot between the BOLD asymmetry index $(BOLD_M - BOLD_E)/(BOLD_M + BOLD_E)$ and corresponding asymmetry index for response times $(RT_M - RT_E)/(RT_M + RT_E)$; (a) left prefrontal and (b) left parietal regions.

1998). However, limitations concerning L2 performance may only apply to particular facets of language processing. For example, ERP patterns in the course of semantic anomaly detection did not appear to vary with age of acquisition of L2, although they did differ for syntactic tasks (Weber-Fox and Neville, 1997). Two studies involving Mandarin–English bilinguals who acquired English relatively late in life did not find significant differences in spatial extent of activation when volunteers generated words in response to partial word-stems or nouns (Chee *et al.*, 1999b; Klein *et al.*, 1999). Additionally, it was found that when language competence was controlled, a later age of acquisition of L2 did not result in a different pattern of activation relative to L1 when volunteers listened to and comprehended sentences (Perani *et al.*, 1998). Finally, it has been shown that usage frequency can compensate for age of acquisition effects (Flege *et al.*, 1999).

Alternatively, the additional activation could be due to an extension of task difficulty and relative proficiency effects. We note that accuracy data obtained during imaging indicated the lowest and slowest performance (in absolute terms) occurred when PRC performed associative judgments in English (their less proficient language). In an experiment involving word semantic associations, the encoding of "distantly linked" word pairs, in comparison to pairings of more closely related words, resulted in greater BOLD signal change in both inferior frontal (and opercular) regions (Fletcher *et al.*, 2000). Similar areas were activated in the present study when English is compared to Mandarin in PRC and could reflect the slight disadvantage experienced by the PRC in performing semantic judgments stemming from differences in their sociocultural background. Activation of the right hemisphere homologues of Broca's and Wernicke's areas was reported when participants read sentences with a more complex structure, compared to those with a less complex structure (Just *et al.*, 1996). While the present study was not designed to elucidate the mechanisms underlying the hypothesized increase in processing demands, it seems reasonable to propose a sequence of activation of brain regions in relation to task difficulty and relative proficiency such that, initially, there is an increase in magnitude of activation in areas already active in performing the task followed by the recruitment of additional brain regions which have the competency to process the task but which are not recruited when task performance is relatively automatic.

Although the proportion of individuals showing left mid and posterior temporal activation did not differ between PRC and SGP, the observation that at the group level of analysis, PRC did not show left mid and posterior temporal activation could indicate differences in functional anatomy between populations. Temporal activations appear to be spatially more varied in PRCs. It has recently been suggested that language background might modulate the topography of brain activation (Paulesu *et al.*, 2000). However, given that temporal activation is noted to be somewhat variable in fMRI studies (Bavelier *et al.*, 1997; Chee *et al.*, 2000; Devlin *et al.*, 2000), it would be prudent to replicate the present finding in a larger cohort before drawing any firm conclusions.

In summary, our findings highlight the importance of taking into consideration the nature of task and relative language proficiency when drawing inferences concerning brain activation in bilinguals processing tasks in different languages.

ACKNOWLEDGMENTS

This work was supported by NMRC Grant 98/00270 and the Shaw Foundation, Singapore. Dr. Alice Seng helped recruit participants for the study. Brendan Weekes provided some helpful comments.

REFERENCES

Bavelier, D., Corina, D., Jezzard, P., Clark, V., Karni, A., Lalwani, A., Rauschecker, J., Braun, A., Turner, R., and Neville, H. 1998. Hemispheric specialization for English and ASL: Left invariance-right variability. *Neuroreport* **9:** 1537–1542.

Bavelier, D., Corina, D., Jezzard, P., Padmanabhan, S., Clark, V. P., Karni, A., Prinster, A., Braun, A., Lalwani, A., Rauschecker, J. P., *et al.* 1997. Sentence reading: A functional MRI study at 4 Tesla. *J. Cogn. Neurosci.* **9:** 664–686.

Chee, M. W., Caplan, D., Soon, C. S., Sriram, N., Tan, E. W., Thiel, T., and Weekes, B. 1999a. Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron* **23:** 127–137.

Chee, M. W. L., Tan, E., and Thiel, T. 1999b. Mandarin and English single word processing studied with fMRI. *J. Neurosci.* **19:** 3050–3056.

Chee, M. W. L., Weekes, B., Lee, K. M., Soon, C. S., Schrieber, A., Hoon, J. J., and Chee, M. 2000. Overlap and dissociation of semantic processing of Chinese characters, English words and pictures: Evidence from fMRI. *NeuroImage* **12:** 392–403.

Chen, P. 1999. *Modern Chinese: History and Sociolinguistics.* Cambridge Univ. Press, Cambridge.

DeFrancis, J. 1984. *The Chinese Language: Fact and Fantasy.* University of Hawaii Press, Honolulu.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van de Moortele, P., Lehericy, S., and Le Bihan, D. 1997. Anatomical variability in the cortical representation of first and second language. *Neuroreport* **8:** 3809–3815.

Demb, J. B., Desmond, J. E., Wagner, A. D., Vaidya, C. J., Glover, G. H., and Gabrieli, J. D. 1995. Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *J. Neurosci.* **15:** 5870–5878.

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., Matthews, P. M., and Tyler, L. K. 2000. Susceptibility-induced loss of signal: Comparing PET and fMRI on a semantic task. *NeuroImage* **11:** 589–600.

Flege, J. E., Yeni-Komishian, G. H., and Liu, S. 1999. Age constraints on Second Language Acquisition. *J. Mem. Lang.* **41:** 78–104.

Fletcher, P. C., Shallice, T., and Dolan, R. J. 2000. "Sculpting the response Space"—An account of left prefrontal activation at encoding. *NeuroImage* **12:** 404–417.

Hernandez, A. E., Martinez, A., and Kohnert, K. 2000. In search of the language switch: An fMRI study of picture naming in Spanish-English bilinguals. *Brain Lang.* **73:** 421–431.

Howard, D., and Patterson, K. 1992. *The Pyramid and Palm Trees Test: A Test of Semantic Access from Words and Pictures.* Thames Valley Test Co., Bury St. Edmunds.

Illes, J., Francis, W. S., Desmond, J., Gabrielli, J. D. E., Glover, G., Poldrack, R., Lee, C. J., and Wagner, A. D. 1999. Convergent cortical representation of semantic processing in bilinguals. *Brain Lang.* **70:** 347–363.

Just, M., Carpenter, P., Keller, T., Eddy, W., and Thulborn, K. 1996. Brain activation modulated by sentence comprehension. *Science* **274:** 114–116.

Kim, K., Relkin, N., Lee, K., and Hirsch, J. 1997. Distinct cortical areas associated with native and second languages. *Nature* **388:** 171–174.

Klein, D., Milner, B., Zatorre, R. J., Zhao, V., and Nikelski, J. 1999. Cerebral organization in bilinguals: A PET study of Chinese-English verb generation. *Neuroreport* **10:** 2841–2846.

Kroll, J. F., and deGroot, A. M. B. 1997. Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In *Tutorials in Bilingualism: Psycholinguisitic perspectives* (J. F. Kroll, and A. M. B. deGroot, Eds.), pp. 169–199. Lawrence Erlbaum Assoc., Malwah.

Moore, C. J., and Price, C. J. 1999. A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* **122:** 943–962.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., Cotelli, M., Cossu, G., Corte, F., Lorusso, M., *et al.* 2000. A cultural effect on brain function. *Nat. Neurosci.* **3:** 91–96.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S., Dupoux, E., Fazio, F., and Mehler, J. 1996. Brain processing of native and foreign languages. *Neuroreport* **7:** 2439–2444.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S., Fazio, F., and Mehler, J. 1998. The bilingual brain—Proficiency and age of acquisition of the second language. *Brain* **121:** 1841–1852.

Rainer, G., and Miller, E. K. 2000. Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* **27:** 179–189.

Rees, G., Friston, K., and Koch, C. 2000. A direct quantitative relationship between the functional properties of human and macaque V5. *Nat. Neurosci.* **3:** 716–723.

Rypma, B., Prabhakaran, V., Desmond, J. E., Glover, G. H., and Gabrieli, J. D. 1999. Load-dependent roles of frontal brain regions in the maintenance of working memory. *Neuroimage* **9:** 216–226.

Smith, E. E., and Jonides, J. 1997. Working memory: A view from neuroimaging. *Cogn. Psychol.* **33:** 5–42.

Talairach, J., and Tournoux, P. 1988. *Coplanar Stereotactic Atlas of the Human Brain.* Thieme Medical, New York.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., and Farah, M. J. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proc. Natl. Acad. Sci. USA* **94:** 14792–14797.

Wang, W.-Y. 1973. The Chinese language. *Sci. Am.* **228:** 50–60.

Weber-Fox, C., and Neville, H. 1997. Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *J. Cogn. Neurosci.* **8:** 231–256.

Xiong, J., Stofer, K. A., Pu, Y., Liu, H. L., Tan, L. H., Gao, J. H., and Fox, P. T. 2000. Different language processing strategy and neural pathways for Chinese speakers. *NeuroImage* **11:** S308.