Mapping of verbal working memory in nonfluent Chinese–English bilinguals with functional MRI

Gui Xue, Qi Dong, Zhen Jin, and Chuansheng Chen

*Institute of Brain and Cognitive Sciences, School of Psychology, Beijing Normal University, Beijing 100875, PR China
Beijing 306 Hospital, Beijing, PR China
Department of Psychology and Social Behavior, University of California, Irvine, CA 92697, USA

Received 30 September 2003; revised 10 January 2004; accepted 12 January 2004

Existing cognitive and neural imaging studies have suggested a frontoparietal network of multiple, cooperative components for verbal working memory (WM). We used functional MRI to investigate whether this neural network is also involved in the processing of second language by nonfluent bilinguals. Twelve (five males, seven females) native Chinese speakers who had limited English proficiency were scanned while performing working memory tasks in Chinese and English. They were asked to make judgment continuously whether the word presented on the screen was semantically related to (i.e., the semantic tasks) another word presented two words earlier. On a different task (i.e., the phonological tasks), they were asked to make judgment whether the target word rhymed with the other word. A naming and judgment task in each language was adopted to control for the visual process, initial lexical process, and motor responses. Behavioral data showed that subjects performed better at tasks in their native language (Chinese, L1) than in English (L2). Imaging results showed that all working memory tasks in both L1 and L2 elicited a very similar pattern of left-hemisphere-dominated activation in the dorsolateral prefrontal cortex, pars opercularis region, precentral cortex, and parietal lobule. Consistent with the behavioral data, the volume of activation in the left opercularis region, left parietal lobule, and right precentral region was greater for L2 than for L1. These results suggest that working memory in L1 and L2 is mediated by a unitary neural system (i.e., frontoparietal region), which is capable of recruiting surrounding cortical resources to meet the increased computational demand caused by low L2 proficiency.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Working memory; Bilinguals; fMRI; Workload; Chinese; Semantic; Phonology

Introduction

Bilingual individuals can perform cognitive functions (e.g., vocabulary acquisition, comprehension, and problem solving) in two languages, but research has not conclusively established whether information processing in the two languages shares the same neural mechanisms. One approach to resolving this issue is to investigate whether the two languages use the same verbal working memory (WM). Verbal WM is the capability to temporally store and manipulate a limited amount of linguistic materials (Baddeley, 1986, 1992), and it is essential to most of the conscious cognitive processes (Jonides, 1995).

Prior cognitive and neural imaging studies have supported a three-component model for verbal working memory: a storage component localized in the left-hemisphere posterior parietal cortex (BA7/40), a subvocal rehearsal component in left-hemisphere speech areas including Broca's area (BA44) and the premotor and supplementary motor areas (BA6), and an executive component in the dorsolateral PFC (BA9/46) (Awh et al., 1996; Honey et al., 2002; Jonides et al., 1997; Paulesu et al., 1993; Smith et al., 1996; see also, Smith and Jonides, 1998, for a review). This three-component model applies to several types of verbal materials, including letters (e.g., Awh et al., 1996; Cohen et al., 1994; Schumacher et al., 1996), digits (Petrides et al., 1993), pseudowords (Barde and Thompson-Schill, 2002), and words (Clark et al., 2000; Crosson et al., 1999). However, it is not known whether WM in second language (L2) involves this cognitive–neural system.

There is clear evidence from cognitive studies with bilinguals that WM processing of L2 relies on the same kind of subvocal rehearsal mechanism as that of L1. First, there is a consistent word-length effect in series recall for both L1 and L2 that can be accounted by articulation rate (Cheung and Kemper, 1993, 1994; Cheung et al., 2000). Second, when concurrent articulation is imposed on a WM task in L2, subjects' performance is dramatically worsened just as is the case in L1 (Cheung et al., 2000). Finally, there is evidence that subjects show similar phonological errors in verbal WM tasks in L2 as in L1 (Osaka and Nishizaki, 2000).

The same cognitive mechanisms for L1 and L2, however, may not mean that the same neural networks are involved. Brain imaging and lesion studies are needed to examine whether L1 and L2 share the same neural mechanisms. So far, the results are inconsistent. On the one hand, research has shown that brain lesion causes equal deficit of working memory capacity in bilinguals' two languages, which suggests common or largely overlapping neural networks for WM processing in L1 and L2 (van Lieshout et al., 1990; Warren et al., 2000). In addition, several functional imaging studies of bilinguals have found common activation for the two languages in the abovementioned three components of the working memory.
system: the executive system (Hernandez et al., 2000, 2001), the speech/rehearsal system (Hasegawa et al., 2002; Kim et al., 1997; Klein, 2003; Klein et al., 1995, 1999; Rodriguez-Fornells et al., 2002), and the retention/storage system (Hasegawa et al., 2002).

On the other hand, a PET study of Korean (L1)–English (L2) bilinguals (Kim et al., 2002) found that the anterior portion of the right dorsolateral prefrontal cortex and the left superior temporal gyrus were activated during WM processing in L1, whereas the posterior portion of the right dorsolateral prefrontal cortex and the left inferior temporal gyrus were activated during WM processing in L2. The authors argued that this result suggested a dissociated neural network for WM processing in L1 and L2. They further argued that the absence of activation for L2 in the subvocal rehearsal and storage system indicated a non-rehearsal-mediated WM processing in L2.

There are at least two explanations for these inconsistent findings. First, different studies have focused on different aspects of WM processing. For example, some studies used tasks that tapped the phonological processes, whereas Kim et al. (2002) used a semantic WM task. As shown by several recent neuroimaging and neuropsychological studies, the semantic and phonological processing involves different areas of frontal cortex (Gabrieli et al., 1998; Poldrack et al., 1999), and the temporary storage of semantic and phonological information relies on separate neurological mechanisms (Martin and Romani, 1994). There is also evidence for differential activation in the semantic processing of bilinguals’ two languages, but not in their phonological processing (Marian et al., 2003). Therefore, it is possible that neural dissociation between L1 and L2 may exist for some aspects of WM processing (e.g., semantic processing), but not for the others (e.g., phonological processing).

A second explanation for the inconsistent results is that the age of acquisition and proficiency level of L2, which varied greatly across studies, may determine the neural mechanisms involved in L2. It has been clearly established that age of exposure has a significant effect on L2 acquisition (Johnson and Newport, 1989, 1991), and may also alter the underlying neural organization for L2 (Kim et al., 1997; Mahendra et al., 2003; Weber-Fox and Neville, 1996, 2001). Cognitive studies also have established a developmental shift from word-mediated to concept-mediated processing as a function of increased L2 proficiency (e.g., Kroll and Sholl, 1992; Kroll and Stewart, 1994). Finally, fMRI and PET studies found significant neural dissociation between L1 and L2 for nonfluent bilinguals (Dehaene et al., 1997; Perani et al., 1996, 1998; Pillai et al., 2003; Wartenburger et al., 2003), but not for highly proficient bilinguals (Chee et al., 1999a,b, 2000; Illes et al., 1999; Klein, 2003; Klein et al., 1995, 1999; Pu et al., 2001). These results suggest that neural dissociation between L1 and L2 may be more likely to occur for late-acquisition (after the age of 12), nonfluent bilinguals than for their early-acquisition, fluent counterparts (see Abutalebi et al., 2001, for a review).

Because of Kim et al. (2002) study’s significance to the argument of neural dissociation for WM processing in L1 and L2, we aimed to replicate their study and to do so with four significant improvements. First, we used a naming and judgment task, instead of a simple picture-detection task, as control. With just a simple picture-detection task as in Kim et al., it is possible that the different activation patterns found for L1 and L2 may result from differences in the initial lexical processing of Korean and English, rather than in WM processing. The impact of surface form of written languages on the reading process and cerebral organization has been demonstrated by several studies involving different language systems, such as Italian vs. English (Paulesu et al., 2000), Chinese vs. English (Liu and Perfetti, 2003), Chinese Pinyin vs. Chinese characters (Chen et al., 2002; Fu et al., 2002), Kanji vs. Kana (Nakamura et al., 2002), English vs. ASL (Neville et al., 1998), and English vs. BSL (MacSweeney et al., 2002). For example, Chinese and English showed the sharpest contrast in surface forms. Accordingly, differential activation was found in the right occipital cortex and the right prefrontal cortex when Chinese–English bilinguals were reading Chinese vs. English words (Liu and Perfetti, 2003). It is thus necessary to minimize the effects of the initial lexical processes when we examine the neural networks underlying the mental storage and manipulation of verbal materials in native and second languages. In our study, a naming and judgment task that involved initial lexical processing was used as a control for this purpose.

Second, we used a typical 2-back paradigm for working memory tasks, whereas Kim et al. (2002) used an atypical 2-back paradigm. In Kim et al. study, participants were asked to monitor the preindicated category (e.g., animal) and make judgment whether a within-category item (e.g., deer) is preceded, with one intervening stimulus, by another within-category item (e.g., tiger). To perform this task, subjects might strategically facilitate their responses by only keeping in mind the sequence of “yes” (i.e., it belongs to the preindicated category) or “no” (i.e., it does not belong to the preindicated category) after each semantic judgment. Consequently, this task may require less temporal maintenance of the lexical information in each language than a typical 2-back paradigm, in which participants had to remember the exact items to make a correct judgment. We believe that a more demanding task involving the typical 2-back paradigm should provide more detection power for neural contrast between WM processing in L1 and L2 than should the atypical 2-back paradigm used by Kim et al.

Third, unlike Kim et al. (2002) study that used only semantic WM tasks, we used two types of tasks in our study: semantic and phonological tasks. In the semantic tasks, subjects were asked to make judgment continuously whether the word presented on the screen was semantically related to another word presented two words earlier. In the phonological tasks, subjects were asked to make judgment whether the target word rhymes with the other word. The inclusion of two types of WM tasks allowed us either to obtain converging evidence across tasks or to determine whether neural dissociation between L1 and L2 exists for some aspects of WM processing (e.g., semantic processing, as found by Kim et al.), but not for the others (e.g., phonological processing) (see Marian et al., 2003, for such a finding).

Finally, our study also dealt with another major confounding variable, namely, memory load. As Hasegawa et al. (2002) pointed out, workload is a major potential confounding variable for studies of nonfluent bilinguals. Many behavioral studies of bilinguals’ working memory have consistently demonstrated that participants would experience a greater difficulty in L2 tasks than in L1 tasks (Cheung and Kemper, 1993, 1994; Cheung et al., 2000; Chincotta and Underwood, 1998; Hoosain, 1979; Stigler et al., 1986; Thorn and Gathercole, 1999, 2001; Thorn et al., 2002). This would inevitably result in differences in neural responses in the WM system because the volume of neural activation in this system is sensitive to computational demand (Braver et al., 1997; Callcott et al., 1999; Cohen et al., 1997; Jonides et al., 1997; Klingberg et al., 1997; Manoach et al., 1997; Rypma and D’Esposito, 1999; Rypma et al., 1999). In this case, a subtraction between the activation maps for L1 and L2 might be inadequate to discover the real pattern of contrast. An alternative method suggested by Hasegawa et al. (2002) is to quantify the volume of activation for each task as
well as the proportion of overlapping activation within each ROIs. This method can provide a quantitative description of the neural dissociation and integration between L1 and L2, as well as an examination of the effect of workload.

In summary, our fMRI study aimed to examine neural integration and dissociation of bilingual WM systems among nonfluent Chinese–English bilinguals. We used a language (Chinese vs. English) by task (semantic vs. phonological WM tasks) factorial design. A naming and judgment task in each language was used to control for the visual process, initial lexical process, and motor responses. Following Hasegawa et al.’s (2002) suggestion, we examined each participant’s data for the extent of overlapping or separate activation in several predefined ROIs involved in verbal WM. To account for the effect of memory load, we also examined differences in volume of activation as a function of language.

Methods

Participants

Fourteen healthy students were recruited from a university campus in Beijing, China. They gave informed consent in accordance with guidelines set by the MRI Center at the Beijing 306 Hospital. Data from two subjects were excluded from analysis because of a malfunction of the fMRI scanner. The remaining 12 subjects (five males and seven females) had a mean age of 20.6 (range 20–24, SD 3.5) years. All subjects were strongly right-handed as judged by the handedness inventory developed by Snyder and Harris (1993).

All subjects completed a questionnaire that asked for their demographic background and self-ratings of English proficiency. All subjects had lived exclusively in China. They began to take English classes when they were 12 years of age or older. All participants were non-English majors, and none of them had had any extracurricular special training in English. Their English skills in reading, speaking, and listening comprehension were self-evaluated on a seven-point scale, where 1 is “not at all skilled” and 7 is “very skilled.” The average ratings were 3.9 for reading (SD = 0.4), 2.8 for speaking (SD = 0.6), and 3.3 for listening comprehension (SD = 0.7). These numbers indicated a low to medium level of proficiency.

Materials and cognitive tasks

Fig. 1 shows the experimental design and examples of the materials used in the study. WM tasks were designed according to the typical 2-back paradigm, in which subjects were asked to make continuous judgment on the semantic (semantically related or not) or phonological (rhyming or not) relations between the word currently presented and the word that had been presented two words earlier in the sequence.1 In the present study, we used a modified definition of a rhyme for English phonological tasks. During the pilot testing, we found that nonfluent bilinguals had trouble differentiating certain consonants (e.g., the sound of th as in teeth vs. the sound of s in rice or glass). Furthermore, these nonfluent bilinguals all pronounced the last consonant as a separate syllable (i.e., rice as ri-seh, glass as gla-seh, teeth as tee-seh) perhaps because all Chinese characters are monosyllable and vowel-ending. Because this is a task designed to ensure that subjects were doing phonological processing of words, we believed that the imaging data should be valid as long as the subjects focused their attention on the pronunciation of words (correctly or incorrectly). Consequently, the definition of a rhyme was modified to decrease the difficulty of English phonological tasks and all subjects were trained according to the modified definition of a rhyme (i.e., “last consonant” rhyming). According to this definition, “teeth,” “glass,” and “rice” rhyme with one another because they all end with the sound of seh for this sample, whereas “apple” and “mirror” do not. As the behavioral data showed (see Fig. 2) the correct ratio (according to the modified definition of a rhyme) for the phonological tasks was high and had no interaction between task and language, indicating that the experimental procedure was likely to be acceptable.

---

1 In the present study, we used a modified definition of a rhyme for English phonological tasks. During the pilot testing, we found that nonfluent bilinguals had trouble differentiating certain consonants (e.g., the sound of th as in teeth vs. the sound of s in rice or glass). Furthermore, these nonfluent bilinguals all pronounced the last consonant as a separate syllable (i.e., rice as ri-seh, glass as gla-seh, teeth as tee-seh) perhaps because all Chinese characters are monosyllable and vowel-ending. Because this is a task designed to ensure that subjects were doing phonological processing of words, we believed that the imaging data should be valid as long as the subjects focused their attention on the pronunciation of words (correctly or incorrectly). Consequently, the definition of a rhyme was modified to decrease the difficulty of English phonological tasks and all subjects were trained according to the modified definition of a rhyme (i.e., “last consonant” rhyming). According to this definition, “teeth,” “glass,” and “rice” rhyme with one another because they all end with the sound of seh for this sample, whereas “apple” and “mirror” do not. As the behavioral data showed (see Fig. 2) the correct ratio (according to the modified definition of a rhyme) for the phonological tasks was high and had no interaction between task and language, indicating that the experimental procedure was likely to be acceptable.
scans, subjects were extensively trained so that they would be familiar with the experimental procedures.

One hundred and twenty Chinese words and 120 English words were selected for the study. Chinese words were all two characters long. The English words had a length of three to seven letters, ranging from one to two syllables (one to three “syllables” as research participants pronounced them, see Footnote 1). All stimulus words were high-frequency nouns with concrete meanings. They were divided into three groups for the semantic WM task, the phonological WM task, and the control task. The use frequency (based on published indices of word use) and length (i.e., number of syllables for English words and number of strokes for Chinese words) of words were carefully matched across the three tasks. To minimize the effect of familiarity, none of the words on the English list had the same meaning as any of the words on the Chinese list.

This experiment involved two scanning sessions, one for Chinese tasks and the other for English tasks. Each run had four semantic blocks, four phonological blocks, and four control blocks. Each block consisted of 10 items and lasted 30 s. The blocks were randomly intermixed. Before each block, there was a cue of 3 s to instruct subjects about the task for that block. Presentation of the two sessions was counterbalanced across subjects. The stimuli were programmed with DMDX on an IBM-compatible notebook and presented by a projector onto a translucent screen. Subjects viewed the stimuli through a mirror attached to the head coil. Materials were presented in white color on black background. For all conditions, each word was presented for 750 ms, with an interstimulus interval (ISI) of 2250 ms. Subjects indicated a positive response by pressing the key corresponding to the index finger of their right hand and a negative response by pressing the key corresponding to the index finger of the left hand.

Apparatus and procedure

Brain scans were performed with a 2.0-T GE/Elscint Prestige whole-body MRI scanner (Elscint Ltd., Haifa, Israel) with standard head coil at the MRI Center of Beijing 306 Hospital. Functional images were obtained by using blood oxygenation level-dependent contrast (Ogawa et al., 1990). The acquisition parameters for single-shot T2*-weighted gradient-echo, EPI sequence was: TR/TE/θ = 3000 ms/60 ms/90°, FOV = 375 × 210 mm, matrix = 128 × 72, and slice thickness = 6 mm. Twenty contiguous axial slices parallel to AC-PC were acquired to cover the whole brain. The anatomical MRI was acquired using a T1-weighted, three-dimensional, gradient-echo pulse-sequence. The parameters for this sequence were: TR/TE/θ = 25 ms/6 ms/28°, FOV = 220 × 220 mm, matrix = 220 × 220, and slice thickness = 2 mm.

Data analysis

Statistical parametric mapping (SPM99, Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab (Mathworks Inc., Sherborn, MA, USA) were used for image preprocessing and statistical analysis. The first three images in each session were excluded from analysis to allow for stability in magnetization. The main steps of image preprocessing included realignment, anatomic–functional image co-registration, spatial normalization (Friston et al., 1995a), and smoothing (8 mm FWHM Gaussian filter). General linear model was used to estimate the condition effect of individual participants (Friston et al., 1995b). Boxcar convolved with HRF was selected as reference function. Individual results were acquired by defining proper effects of interests for each subject with the relevant parameter estimates. The threshold for significant activation was P < 0.05 (multiple-comparison corrected). The group-averaged effects were computed with a random-effects model. For group results, clusters with more than 10 voxels (3 × 3 × 3 mm) activated above a threshold of P < 0.0001 (uncorrected) were considered as significant.

ROIs selection and quantitative analysis

To quantify the extent of neural integration and dissociation and to compare the amount of activation in a given area across experimental conditions, several brain areas associated with reliable activation in verbal WM tasks and their right homologous
**Results**

**Behavioral data**

As Fig. 2 shows, there were significant main effects of both language \( F(1,11) = 29.7, P < 0.0002 \) and task \( F(1,11) = 6.102, P < 0.031 \], but no significant interaction between them \( F(1,11) = 0.14, n.s. \). More errors were made for the English tasks than for Chinese WM tasks in the left pars opercularis region (Fig. 4c). Marginally significant language effects were also found in the left parietal lobule (Fig. 4e) and right precentral region (Fig. 4g). No region was more extensively activated for Chinese tasks than for English tasks. The phonological tasks elicited significantly more activation than did the semantic tasks in left pars opercularis region (Fig. 4c), left

<table>
<thead>
<tr>
<th>Region of interest</th>
<th>Semantic tasks (%)</th>
<th>Phonological tasks (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left dorsolateral PFC (Fig. 4a)</td>
<td>63</td>
<td>61</td>
<td>62</td>
</tr>
<tr>
<td>Left precentral cortex (Fig. 4b)</td>
<td>77</td>
<td>80</td>
<td>78</td>
</tr>
<tr>
<td>Left pars opercularis (Fig. 4c)</td>
<td>79</td>
<td>79</td>
<td>76</td>
</tr>
<tr>
<td>Left pars triangularis (Fig. 4d)</td>
<td>68</td>
<td>74</td>
<td>71</td>
</tr>
<tr>
<td>Left parietal lobule (Fig. 4e)</td>
<td>53</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>Left precentral cortex (Fig. 4f)</td>
<td>62</td>
<td>57</td>
<td>59</td>
</tr>
<tr>
<td>Left pars opercularis (Fig. 4g)</td>
<td>32</td>
<td>45</td>
<td>38</td>
</tr>
<tr>
<td>Left pars triangularis (Fig. 4h)</td>
<td>52</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>Right parietal lobule (Fig. 4i)</td>
<td>51</td>
<td>58</td>
<td>55</td>
</tr>
</tbody>
</table>

were selected as ROIs. They were the dorsolateral prefrontal ROIs corresponding to the mid-dorsal frontal gyrus (BA9/46), the par opercularis ROIs in the posterior portion of the ventrolateral prefrontal gyrus (BA44), the par triangularis ROIs in the anterior portion of the ventrolateral prefrontal gyrus (BA45/47), the precentral ROIs including the premotor cortex and SMA (BA 6), the parietal ROIs including the superior parietal lobule (BA7), inferior parietal lobule (BA40), supramarginal gyrus (BA39), and angular gyrus (BA40).

To assess the extent of overlapping activation, we calculated the proportion of voxels activated in both languages compared to the averaged total number of activated voxels in Chinese and English. In addition, we counted the number of significantly activated voxels in each ROI for all four language × task conditions and conducted ANOVAs with repeated measures to evaluate the main effects of language and WM task and their interaction.

**Overlapping activation**

We quantified the extent of overlapping activation across languages and tasks (see Table 1). For cross-language overlap in the semantic WM tasks, for example, we counted for each participant the number of activated voxel within each ROI for Chinese semantic WM tasks (CS), English semantic WM tasks (ES), the commonly activated voxels in the two tasks (CES), then computed the proportion of overlapping by the following formula, common activation% \( = \frac{2 \times CES}{CS + ES} \times 100\% \).

On average, we got a high degree of cross-language overlap (greater than 60%) for the semantic and phonological tasks in left-side ROIs. This high degree of overlap is especially noteworthy when one considers the stability of fMRI measurement of neural activation (Aguirre et al., 1998). These results strongly suggest that a highly overlapping neural network mediates the WM tasks in both Chinese and English. The relatively lower degree of overlap in the right-side ROIs might be due to the overall decreased volume of activation in these regions (see Fig. 3).

**Volume of activation**

We conducted two-way ANOVA on the number of activated voxels for each ROI (see Table 2). There was a significantly greater amount of activation for English than for Chinese WM tasks in the left pars opercularis region (Fig. 4c). Marginally significant language effects were also found in the left parietal lobule (Fig. 4e) and right precentral region (Fig. 4g). No region was more extensively activated for Chinese tasks than for English tasks. The phonological tasks elicited significantly more activation than did the semantic tasks in left pars opercularis region (Fig. 4c), left

**Table 1**

<table>
<thead>
<tr>
<th>Region of interest</th>
<th>Semantic tasks (%)</th>
<th>Phonological tasks (%)</th>
<th>Total (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left dorsolateral PFC (Fig. 4a)</td>
<td>63</td>
<td>61</td>
<td>62</td>
</tr>
<tr>
<td>Left precentral cortex (Fig. 4b)</td>
<td>77</td>
<td>80</td>
<td>78</td>
</tr>
<tr>
<td>Left pars opercularis (Fig. 4c)</td>
<td>79</td>
<td>79</td>
<td>76</td>
</tr>
<tr>
<td>Left pars triangularis (Fig. 4d)</td>
<td>68</td>
<td>74</td>
<td>71</td>
</tr>
<tr>
<td>Left parietal lobule (Fig. 4e)</td>
<td>53</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>Left precentral cortex (Fig. 4f)</td>
<td>62</td>
<td>57</td>
<td>59</td>
</tr>
<tr>
<td>Left pars opercularis (Fig. 4g)</td>
<td>32</td>
<td>45</td>
<td>38</td>
</tr>
<tr>
<td>Left pars triangularis (Fig. 4h)</td>
<td>52</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>Right parietal lobule (Fig. 4i)</td>
<td>51</td>
<td>58</td>
<td>55</td>
</tr>
</tbody>
</table>

\( a: \) marginal effect.

* Significant effect \( (P < 0.05) \).
precentral region (Fig. 4b), left parietal region (Fig. 4e), and right parietal region (Fig. 4j). A marginally greater amount of activation for the phonological tasks than for the semantic tasks was also detected in the right pars opercularis region (Fig. 4h). No region was more extensively activated for the semantic tasks than for the phonological tasks. The bilateral dorsolateral PFC region (Figs. 4a, f) and bilateral left pars triangularis region (Figs. 4d, i) were similarly activated across all conditions.

Discussion

In this section, we first discuss the validity of our experimental manipulations in reference to the existing cognitive–neural model. We then discuss our findings regarding the neural integration and dissociation of L1 and L2 for the semantic and phonological WM tasks. Finally, we discuss the role of memory load in modulating the activation of verbal WM network.

Neural network for verbal WM processing

Using semantic and phonological 2-back paradigm, the present study found that our WM tasks resulted in significant activation in the Broca’s area, the premotor and the SMA, and the parietal area. This profile of activation has been well replicated in previous studies using a similar paradigm or the item-recognition paradigm employing letters, digits, and words (e.g., Clark et al., 2000; Cohen et al., 1997; Fiez et al., 1996; Honey et al., 2002; Jonides et al., 1997; Paulus et al., 1993; Petrides et al., 1993; Smith et al., 1996). These results suggest that common subvocal rehearsal and storage processes underlie all verbal WM processes (Smith and Jonides, 1998).

Consistent with previous studies (Cohen et al., 1997; Petrides et al., 1993; Postle et al., 1999; Rowe et al., 2000; Wagner et al., 2001), we also found that the dorsal lateral prefrontal cortex (DLPFC) was activated. This can be because, compared to the control tasks (naming and judgment), our WM tasks required extensive cognitive resource to update and check WM content and to code the temporal sequence of these stimuli. Furthermore, we found that the bilateral DLPFC was almost equivalently activated, suggesting that the executive function may be less lateralized than previously believed.

Our study also found activation in the pars triangularis area, an anterior portion of ventrolateral prefrontal cortex (aVLPFC). Although such activation was not commonly reported in previous studies with letters and numbers, our results are consistent with several studies that, like our study, used words as stimuli (Clark et al., 2000; Crosson et al., 1999; Thierry et al., 2003). Taken together, these results seem to suggest that the aVLPFC is crucial for the processing of words, but not letters or numbers.

In sum, there is evidence that our experimental manipulations of tasks produced imaging results that are comparable to those of previous studies of verbal WM. Next, we discuss neural dissociation and integration between L1 and L2 for working memory tasks.

Neural dissociation and overlap between L1 and L2 in WM

The major finding of the present study is that our bilingual subjects showed highly overlapping brain activation when processing L1 and L2. Significant activations were consistently found in semantic and phonological WM for the two languages in the bilateral DLPFC, precentral cortex, pars opercularis cortex, pars triangularis cortex, and parietal lobule. Within-ROI computation of...
overlapping activation between L1 and L2 revealed a considerable proportion of co-activation in several regions crucial for verbal WM processing, including the left DLPFC, left precentral cortex, left pars opercularis region, left pars opercularis region, and left parietal lobule. Thus, our results provide compelling evidence that the WM processing of L1 and L2 share the same neural correlates.

In line with behavioral data on bilinguals’ working memory (see Introduction), the extensive activation in the left pars opercularis region, the precentral region, and the parietal lobule found in the present study for L2 suggested a similar subvocal rehearsal and storage mechanism underlying L2 WM processing. This result does not support the hypothesis that a visual strategy would necessarily be adopted to compensate for the low L2 proficiency (Kim et al., 2002). In addition, the high degree of overlapping activation observed in the DLPFC, an area linked to the executive component of WM, provided evidence that the manipulation of verbal materials in L1 and L2 shared the same neural mechanism.

The absence of a significant language by task interaction provides convergent evidence to support a common neural network for WM processing in L1 and L2. It should be noted that the experimental paradigm adopted in the present study does not allow us to have a strict control over how words are stored in the working memory. Although there is evidence in brain lesion patients that semantic information is stored separately from phonological information (Martin and Romani, 1994), it is still possible that, in a 2-back semantic judgment task, subjects could store each word phonologically only and wait until the time of judgment to retrieve the semantic information. Actually, we found very similar activation pattern for semantic and phonological WM tasks, which might be due to the similar neural substrates for the two processes (e.g., Barde and Thompson-Schill, 2002), or the compulsive co-activation of semantic and phonological information during word processing (MacLeod, 1991), or a common rehearsal/storage strategy as suggested above.

Because our findings of similar neural mechanisms for L1 and L2 WM processing came from nonfluent bilinguals who acquired their second language at a late age, it seems likely that there is not a general critical period for the neural organization of L1 and L2. Instead, such a critical period, if it exists, may be task-specific. For example, prior research showed dissociated activation for late bilinguals in sentence-generation task (Kim et al., 1997; Mahendra et al., 2003) and word fluency task (Mahendra et al., 2003), but not for story comprehension task (Perani et al., 1998) and semantic decision task at single-word level (Chee et al., 1999b; Illes et al., 1999). Task-specific dissociation was also found in studies that directly compared semantic and syntactic processing between bilinguals who acquired L2 at an early age and those at a late age (Wartenburger et al., 2003; Weber-Fox and Neville, 1996, 2001; also see Fabbro, 2001, for a review). It should be noted, however, that some of the empirical evidence for neural dissociation for late-acquisition bilinguals might have been confounded by language proficiency (Perani et al., 1998), time of exposure (Perani et al., 2003), or task difficulty (Chee et al., 2001). To disentangle these confounding factors (especially task difficulty), the effect of workload needs to be taken into consideration (Hasegawa et al., 2002), a point to which we will return in the next section.

Our results appear to contradict the Kim et al. (2002) study that revealed neural disassociation for WM processing in native and second languages. Given the different findings, it is certainly premature to draw any definitive conclusions. For the sake of argument, however, we can discuss several other aspects of Kim et al. study (in addition to the ones mentioned in the Introduction) that may present further challenges to their conclusion. First, in their study, the argument for a visual strategy for WM processing in English was partially based on the similar pattern of activation for English and simple pictures. However, it was possible for the subjects to use a verbal strategy for the picture tasks. Second, existing functional imaging studies do not seem to support Kim et al.’s claim of the anterior/posterior dissociation in right dorsolateral PFC in mediating the manipulation of verbal and visual materials. Instead, the evidence is for a left/right dissociation in mediating the manipulation of verbal and visual materials (Smith et al., 1996), or bilateral activation for both verbal and nonverbal tasks (Nystrom et al., 2000; also see D’Esposito et al., 1998, for a review). Third, due to insufficient control for lexical processes in Kim et al. study as mentioned before, it is not clear whether the differential activation in the temporal region reflects distinct working memory processes, or rather different routes for lexical processes (Jobard et al., 2003; Price, 2000). Finally, the absence of action in several regions relevant to WM in Kim et al.’s study might be partly due to the low statistical detection power and/or the atypical 2-back design’s decreased demand on working memory (see Introduction).

Putting together our findings with the evidence of neural contrast between subjects’ two languages during reading (e.g., Chen et al., 2002; Liu and Perfetti, 2003), we may speculate that lexical information is temporarily stored and manipulated in the same working memory system, but the processing (mainly the visual identification, lexical—semantic, and lexical—phonological processing) of the two languages may use different neural (and/or cognitive) mechanisms due to factors such as the design principle of different language systems. We hasten to add, however, the usefulness of this speculation is limited by the low spatial and temporal resolution of fMRI method that may have failed to detect subtle differences in the cognitive and neural operations of working memory for the two languages. Methods with better resolution such as functional magnetic resonance adaptation (Chee et al., 2003), or combination of ERPs and fMRI, should help to clarify this issue.

Workload and volume of activation in bilingual WM processing

Finally, we should comment on the effects of workload on the brain activation. Prior research has shown workload effects involving tasks such as sentence reading (Just et al., 1996; Keller et al., 2001), word-span task (Grasby et al., 1994), mental rotation task (Carpenter et al., 1999), and word reading task (Price et al., 1992). Specifically, many studies suggested that the WM networks are load-sensitive. When item load increases systematically (e.g., from 0-back to 3-back), there is a general increase in functional activation in the frontal—parietal cortical regions (Barch et al., 1997; Cohen et al., 1997; Jonides et al., 1997; Manoach et al., 1997; Rypma et al., 1999), although the load—response curve was not always linear (Cohen et al., 1997; Jonides et al., 1997). Studies on nonfluent bilinguals also found workload differences between their native and second languages (Chee et al., 2001; Hasegawa et al., 2002). It seems that the workload effect would explain our subjects’ significantly poorer performance in L2 than in L1 for both semantic and phonological tasks. In trying to compensate for L2 low proficiency, especially that resulted from difficulty in articulating a L2 word (Cheung and Kemper, 1993, 1994; Cheung et al., 2000) or a limited knowledge structure of L2 (Thorn and Gathercole, 1999; Thorn et al., 2002), subjects might passively
and/or strategically have put more resources on rehearsal in L2. This speculation is consistent with the greater activation in the Broca’s area, the left parietal lobule, and the right precentral cortex for L2 than for L1 tasks.

On the other hand, the present study did not find any workload effect in the executive region including bilateral DLPFC for the two languages. To date, parametric studies of working memory have obtained inconsistent results. For example, in studies by Cohen et al. (1997) and Jonides et al. (1997), the BOLD signal change was not increased linearly as the n increase from 0 to 3. Increase in activation with item number did not appear until moving to 2-back task. There was no significant difference between 0-back and 1-back tasks and between 2-back and 3-back tasks. Another study with a similar design found a roughly linear increase in DLPFC with memory load (Braver et al., 1997). Further investigation is needed to explore the computational nature of executive function.

We also found that the poorer behavioral performance in the phonological tasks than in the semantic tasks was accompanied by increased activation in the subvocal rehearsal and storage system, namely, the bilateral pars opercularis region, left precentral cortex, and bilateral parietal cortex. One possibility for this semantic advantage is that the semantic information might be accessed via a non-phonological way (Coltheart et al., 1993; Fiebach et al., 2002). Thus, semantic decision can be made without pronouncing the currently presented word (though articulation is necessary to keep this word for the following task). This could greatly decrease the workload of phonological rehearsal. This explanation, however, is highly speculative. It is entirely possible that the task-specific activation that appears to parallel task difficulty was not caused by workload, but rather by task-specific requirement that increased activation in these brain regions (Demb et al., 1995; also see Chee et al., 2001, for a discussion). Future research should use multiple levels of task difficulty within each task to examine the workload effects.

In summary, this study examined the neural substrates underlying WM processing in L1 (Chinese) and L2 (English). The considerable overlapping activations for L1 and L2 in the fronto-parietal regions (including DLPFC, pars opercularis region, pars triangularis region, precentral cortex, and parietal lobule) suggest a very similar neural system for bilingual WM processing. Quantitative analysis further suggests that this neural system is modulated by a compensatory and/or passive mechanism of recruiting surrounding cortical resource to meet the increased computational command caused by low L2 proficiency.

Acknowledgments

This work was supported by national Pan Deng project (95-Special-09). We would also like to thank two anonymous reviewers for their helpful comments.

References


