Processing of Visually Presented Sentences in Mandarin and English Studied with fMRI

Michael W. L. Chee,†∥ David Caplan,‡ Chun Siong Soon,† N. Sriram,† Edsel W. L. Tan,* Thorsten Thiel,§ and Brendan Weekes†
†Cognitive Neuroscience Laboratory
Singapore General Hospital
Singapore 169856
‡Department of Social Work and Psychology
National University of Singapore
Singapore 119260
§Neuropsychology Laboratory
Department of Neurology
Massachusetts General Hospital
Boston, Massachusetts 02114
∥Department of Radiology
University of Freiburg
D79106 Freiburg
Federal Republic of Germany
³Neuropsychology Laboratory
Department of Psychology
University of Kent
Canterbury CT27LZ
United Kingdom

Summary

Comprehension of visually presented sentences in fluent bilinguals was studied with functional magnetic resonance imaging (fMRI) using a set of conceptually similar sentences in two orthographically and phonologically distinct languages, Mandarin and English. Responses were monitored during scanning. Sentence comprehension in each language was compared to fixation in nine subjects and Tamil-like pseudo-word strings in five subjects. Spatially congruent activations in the prefrontal, temporal, and superior parietal regions and in the anterior supplementary motor area were observed for both languages and in both experiments at the individual and group levels of analysis. Proficient bilinguals exposed to both languages early in life utilize common neuroanatomical regions during the conceptual and syntactic processing of written language irrespective of their differences in surface features.

Introduction

The goal of this study was to elucidate the functional anatomy of Mandarin and English written sentence comprehension in bilinguals using functional magnetic resonance imaging (fMRI). Specifically, we wanted to determine if differences in the surface features (orthography, phonology, and syntax) of different languages affect their cerebral organization at the sentence level of processing.

Separate lexicons (repositories for information about words) for the two languages have been demonstrated by a variety of behavioral experiments (reviewed by Smith, 1997). However, it remains unclear whether these lexicons are spatially segregated within the brain. The case for anatomically segregated lexicons is suggested by clinical case studies of individuals who experience dissociated language loss following brain damage (Albert and Obler, 1978) as well some data from electrical stimulation of the brain (Ojemann and Whitaker, 1978; Rapport et al., 1983). An alternative explanation for the loss or preferential recovery of one language in a bilingual postulates disruption of a control system that allocates resources to one or the other language (Green, 1986; Paradis, 1998). Under this framework, spatially segregated lexicons are not required to explain dissociated language loss or recovery.

Neuroimaging studies of language function in healthy bilingual subjects have sought to demonstrate neural substrates responsible for bilingual language processing. However, the work to date has yielded conflicting results. At the single word level, several studies report overlapping activations in two languages (Klein et al., 1994, 1995; Chee et al., 1999), whereas, at the sentence level, intrahemispheric differences in left hemisphere activations as well as occasional interhemispheric differences have been reported (Mazoyer et al., 1993; Perani et al., 1996; Dehaene et al., 1997; Kim et al., 1997). One explanation for the contrasting results is that, in comparison to word level processing, sentence level processing requires syntactic construction and additional working memory resources that differ across languages. Alternatively, this result may be consequent to lack of adequate control of parameters that affect topography of activation but that are not directly related to structural differences between the languages.

Several studies have used sentence stimuli based on translations of a story (Mazoyer et al., 1993; Perani et al., 1996; Dehaene et al., 1997). In these studies, subjects answered questions relating to a passage that was read to them, only after the imaging session. In an fMRI study requiring sentence production, subjects recalled what they did at various times of the day (Kim et al., 1997). In each of these experiments, subjects’ responses during scanning were not monitored, and it is uncertain what cognitive processes were engaged during the experiment. Specifically, there were no overt attempts to reduce the use of translation from L2 to L1 or code switching during the experiment. A second factor that has not been consistently controlled in previous studies is the relative fluency of the subjects in the second language (L2).

In this study, we presented sentences visually to fluent English–Mandarin bilinguals, probed for comprehension following each presentation (Just et al., 1996; Keller et al., 1998, Neuroimage, abstract), and analyzed individuals’ activations in both languages. We presented English and Mandarin sentences because these languages differ in orthography, phonology, and syntax. By selecting proficient bilinguals exposed to both languages early in life, we sought to minimize the effects of age of acquisition of L2 (Weber-Fox and Neville, 1997; Neville and Bavelier, 1998) and language proficiency (Perani et al., 1998) so that any topographical differences in activation...
Table 1. Accuracy and Reaction Time Data

<table>
<thead>
<tr>
<th></th>
<th>Accuracy (%) Correct</th>
<th>Reaction Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Behavioral pilot (n = 15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>88.4</td>
<td>5.0</td>
</tr>
<tr>
<td>Mandarin</td>
<td>89.2</td>
<td>5.5</td>
</tr>
<tr>
<td>fMRI Experiment 1 (n = 9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>89.6</td>
<td>5.5</td>
</tr>
<tr>
<td>Mandarin</td>
<td>88.4</td>
<td>6.3</td>
</tr>
<tr>
<td>fMRI Experiment 2 (n = 5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>89.7</td>
<td>10.5</td>
</tr>
<tr>
<td>Mandarin</td>
<td>87.0</td>
<td>10.4</td>
</tr>
</tbody>
</table>

well as to control for motor processes involved in responding to the stimuli.

Behavioral Data

Accuracy and reaction time data were analyzed using SPSS 8.0 (SPSS, Chicago, Illinois) and Statview 4.5 (Abacus Concepts, Berkeley, California). Language (English or Mandarin) and clause (first or second) were the two factors used in the analysis. Means and standard deviations for reaction times and accuracy for each experiment are shown in Table 1.

In the pilot behavioral study, there was no difference between performance in English and Mandarin in either reaction time (F[1,14] = 2.35, p = 0.15, NS) or accuracy data (F[1,14] < 1, NS) as shown in Table 1. Reaction times were longer when the first clause was probed (F[1,14] = 21.8, p < 0.0001). The interactions between language and clause were not significant (both Fs < 1). These results were replicated in the behavioral data gathered during both fMRI experiments. There was no difference between performance in English and Mandarin in either reaction time (F[1,8] < 1, NS; F[1,4] < 1, NS) or accuracy data (F[1,8] < 1, NS; F[1,4] < 1, NS).

Activations Occurring during Sentence Interpretation versus Fixation

Compared to fixation, regions activated by sentence interpretation were: inferior (BA 44, 45, 47) and middle prefrontal cortex (BA 9, part of BA 8, BA 6; more extensive on the left side), left temporal region (BA 22, 21, 38), left angular gyrus (BA 39), anterior supplementary motor area (SMA) (BA8; left-sided activations predominating), and bilateral superior parietal (BA 7) and occipital regions (Table 2; Figures 2 and 3).

Frontal activations were robust. All subjects showed activations in the superior portions of BA 44, 45 in the upper part of the inferior frontal gyrus or the middle frontal gyrus (BA 9).

Temporal lobe activations were seen in seven out of nine subjects. Six subjects activated the left posterior superior temporal gyrus and/or the middle temporal gyrus with extension into the supramarginal or angular gyri. In addition, three subjects showed activations in the anterior temporal pole. Left inferior temporal regions were revealed in five subjects. Prefrontal activations were more extensive than temporal activations but the latter were more biased toward the left hemisphere.

The use of a lower statistical threshold (Z score, 3.5
instead of 5.0) showed an increase in extent of activated areas (Figures 2 and 5A). This effect was more pronounced in the right hemisphere and caused a shift in the asymmetry index (AI) (Table 3).

Activations Occurring during Sentence Interpretation versus Tamil-like Strings

As expected from controlling low-level perceptual processes, occipital activations were reduced but not completely suppressed. With minor variations, activations of the frontal and temporal areas were similar in spatial location to those described in the first experiment (Figures 4 and 5A). Experiment 2 was associated with greater left hemisphere bias of frontal and temporal activations compared to Experiment 1 (Table 3). The pooled data had a frontal AI of 0.97 for Mandarin and 0.99 for English at a Z score threshold of 5.0. The corresponding AIs for Experiment 1 were 0.73 and 0.67.

Extent of Overlap of Activations and Direct Comparison between English and Mandarin

At both individual and pooled data levels, activations in the prefrontal and temporal regions overlapped extensively during English and Mandarin sentence comprehension in both experiments (Tables 2 and 3; Figures 2-5A). Direct comparison of English and Mandarin yielded no active pixels even when subtle differences in the location or extent of activation appeared in the language versus fixation or language versus pseudoword control conditions. The frontal AIs between English and Mandarin were highly correlated across subjects (Pearson correlation coefficient = 0.85, Experiment 1) and across both languages.

Our initial analysis was based on a Z score threshold of 5 in order to minimize false positives (fortuitous, non-task-locked variations in MR signal; “rim” artifacts from motion; activations within ghosts). Some of these artifacts were seen at the Z score 3.5 level (Figure 4; inferior slices in HT and AV show ghost artifacts anterior to the frontal regions). To deal with concerns that subtle differences in activation between languages might be obscured, maps with the lower threshold are shown. Importantly, lowering the detection threshold to 3.5 did not uncover significant intrahemispheric or interhemispheric differences in activations between the languages.

Effect of Pooling of Baseline Epochs

Fixation periods were pooled with a view to increase the number of signal samples related to this baseline period. To allay concerns that signals from the fixation period following each language task are not equivalent, we compared activation maps for English versus fixation under three conditions (fixation following English, fixation following Mandarin, and pooled fixation; Figure 5B).
As illustrated, the signals obtained from different epochs of fixation do not result in significant differences in activations.

Discussion

This study demonstrates that in fluent bilinguals, common areas are activated during the comprehension of syntactically complex English and Mandarin sentences, supporting a "one-store" model for the linguistic representation of two languages.

Much of the impetus for supporting a "two store" model comes from clinical case reports of dissociated language loss and recovery (Albert and Obler, 1978; Paradis, 1995). While this suggests differential spatial localization of the first (L1) and second (L2) language, a few points should be noted. Firstly, these individuals represent a reportable minority of aphasic patients who suffer language loss (Penfield and Roberts, 1959). Secondly, a two store model does not explain why some bilingual aphasics have access to only one language for alternating periods of time (Paradis, 1998) (see also the discussion below). Relevant to the comparison of different orthographies, recent studies of Kanji (ideographic script) versus Kana (phonographic script) processing using event-related potential (ERP) (Koyama et al., 1998) and lesion studies (Sugishita et al., 1992) did not reveal the traditional functional-anatomical dissociations between the two scripts.

Differences in the locations at which cortical stimulation interferes with naming in the two languages of bilingual subjects have been another source of support for the two store model (Ojemann and Whitaker, 1978). In Ojemann's study, neither subject was a neurologically intact individual. One subject had a prior left temporal lobectomy and a low verbal IQ. The other was left-handed with right hemisphere language dominance. Another electrical stimulation study specifically examining Chinese-English polyglots also showed a few stimulation sites where language was spatially segregated (Rapport et al., 1983). With both studies, sites where stimulation interfered with naming in both languages outnumbered those where one language was affected. Sites where there was a clear double dissociation of naming interruption were rare.

As languages can be represented across phonological, orthographic, semantic, pragmatic, and discourse dimensions, a thorough investigation into the functional anatomy of language in bilinguals will have to look at each of these dimensions following the hierarchy of linguistic complexity. In contrast to the numerous studies on single word processing, there have been relatively...
Figure 3. Individual and Pooled Activation Maps from Experiment One
Montage of activations of all nine subjects participating in Experiment One (where fixation was the control condition). The subject's left hemisphere is on the right side of each image. Areas activated above a Z score of 3.5 are shown. Green areas indicate an overlap of English (blue) and Mandarin (yellow) activations. The images are oriented in Talairach space. Subjects' initials appear on the upper right hand corner of each strip. “AV” represents the pooled data set; shown in this strip are (A) the anterior temporal pole (BA 38), (B) the mid temporal region (BA 21, 22), (C) the posterior superior temporal gyrus (BA 22), (D) the prefrontal region (BA 44, 9), (E) the anterior SMA, and (F) the superior parietal region (BA 7).

Few imaging studies characterizing sentence processing (Mazoyer et al., 1993; Just et al., 1996; Perani et al., 1996, 1998; Stromswold et al., 1996; Bavelier et al., 1997; Dehaene et al., 1997; Müller et al., 1997; Caplan et al., 1998, Neuroimage, abstract; Dapretto et al., 1998, Neuroimage, abstract; Keller et al., 1998, Neuroimage, abstract). Studying language at this level has the advantage of using a cognitively natural task (Poeppel, 1996; Bavelier et al., 1997). While we may be criticized for not attempting to relate subcomponents of sentence comprehension (Osterhout et al., 1997) to brain regions, our strategy is to cast a wide net in search of potential functional-anatomical differences prior to evaluating specific processes.

Existing work suggests that analysis of visual word form recruits left extrastriate and inferior temporal areas (Nobre et al., 1994). Syntactic and sentential semantic analyses activate the inferior frontal (Just et al., 1996; Stromswold et al., 1996; Caplan et al., 1998, Neuroimage, abstract; Gabrieli et al., 1998), anterior temporal (Bavelier et al., 1997), and posterior temporal areas (Just et al., 1996; Keller et al., 1998, Neuroimage, abstract).

Verbal working memory activates a network of areas that includes the dorsolateral prefrontal cortex (BA 9, 46), the middle frontal gyrus and premotor area (BA 6), the anterior SMA, the posterior parietal cortex (BA 7), and the right cerebellar hemisphere (Fiez et al., 1996; Kelley et al., 1998; Smith et al., 1998). Some of these assignments of linguistic functions to specific brain regions should be regarded as preliminary, as different regions seem to be highlighted depending on the experimental design utilized.

A less modular view of the organization of sentence processing resources emerges from a study of patients with syntactic processing deficits following stroke. While left perisylvian lesions were overrepresented in these patients, there was no difference between the performance of patients with anterior and posterior lesions and no correlation between the degree of impairment and the size of lesions (Caplan et al., 1996). This result favors a distributed neural net model (McClelland et al., 1989) for sentence processing with a bias toward regional specialization of left perisylvian neural circuits for such processing. Under this schema, the size of the
Table 3. Voxel Counts of Frontal and Temporal Activations in Individual and Pooled Data Sets

<table>
<thead>
<tr>
<th>Frontal Activations (voxels)</th>
<th>Temporal Activations (voxels)</th>
<th>Asymmetry Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandarin</td>
<td>English</td>
<td>Mandarin</td>
</tr>
<tr>
<td>R</td>
<td>L</td>
<td>R</td>
</tr>
<tr>
<td>Experiment 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GR</td>
<td>5</td>
<td>488</td>
</tr>
<tr>
<td>KC</td>
<td>103</td>
<td>570</td>
</tr>
<tr>
<td>LI</td>
<td>152</td>
<td>691</td>
</tr>
<tr>
<td>SH</td>
<td>92</td>
<td>156</td>
</tr>
<tr>
<td>TJ</td>
<td>298</td>
<td>863</td>
</tr>
<tr>
<td>YE</td>
<td>36</td>
<td>731</td>
</tr>
<tr>
<td>YU</td>
<td>147</td>
<td>710</td>
</tr>
<tr>
<td>ZI</td>
<td>70</td>
<td>181</td>
</tr>
<tr>
<td>YO</td>
<td>179</td>
<td>520</td>
</tr>
<tr>
<td>AV3.5</td>
<td>48</td>
<td>46</td>
</tr>
<tr>
<td>AV5.0</td>
<td>73</td>
<td>67</td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH</td>
<td>20</td>
<td>42</td>
</tr>
<tr>
<td>PL</td>
<td>6</td>
<td>222</td>
</tr>
<tr>
<td>HT</td>
<td>27</td>
<td>267</td>
</tr>
<tr>
<td>LI</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>DE</td>
<td>46</td>
<td>380</td>
</tr>
<tr>
<td>AV3.5</td>
<td>73</td>
<td>88</td>
</tr>
<tr>
<td>AV5.0</td>
<td>97</td>
<td>99</td>
</tr>
</tbody>
</table>

Voxels above the threshold of Z score $\geq 5$ were counted for individual subjects. AV3.5 and AV5.0 refer to the pooled data sets from each experiment where Z score thresholds of 3.5 and 5.0 were used. The asymmetry index was computed for the frontal regions only: $AI = \frac{\text{sum} \{\text{voxels (L - R))\}}{\text{sum} \{\text{voxels (L + R))\}}$. In Experiment 1, the Pearson correlation coefficient between Mandarin and English AI was 0.85.

network recruited to process sentences increases with task complexity (Just et al., 1996).

In Experiment 1, where fixation was the control, we had intended to uncover differences in the early phases of letter or character recognition. We did not find differences, perhaps because of the choice of imaging plane. Alternatively, word or word-like stimuli may activate the inferior temporal region irrespective of orthography as suggested by a magnetoencephalography (MEG) study of Kanji and Kana (Koyama et al., 1998). In Experiment 2, we sought to control for perceptual processes related to character or word recognition and motor aspects of responding to the probe question. The location of prefrontal and temporal activations was comparable to those seen in the first experiment. The results of both sets of experiments suggest that the entire gamut of processes required for comprehending visually presented sentences in English and Mandarin activates a common set of brain regions in a given individual.

Our findings extend those of Perani (Perani et al., 1998) who found overlapping activations in response to attending to auditory sentences in Spanish and Catalan among fluent bilinguals. The Romance languages compared in Perani’s study have identical orthography, relatively similar phonology, and a high degree of overlap at the lexical and syntactic levels. This makes the interpretation of their data difficult in terms of whether the overlapping regions reflect shared conceptual processing or the processing of common surface features.

Mandarin has a completely different orthography and phonology to English and hence the two languages do not share cognates, words with similar sounds and meanings across languages. There is no inflectional morphology in Mandarin so that syntactic cues are relatively few. This arguably makes sentence processing more dependent on context and working memory compared to English. The task we used employed conceptually similar stimuli across languages. Finally, online monitoring of performance was carried out to ensure that the subject performed the task in the desired language. Taken together, these points allow us to argue more convincingly that the areas of common brain activation seen in the present study reflect common conceptual processing across languages rather than processing of shared surface features. The use of fMRI allowed the study of intraindividual activations as distinct from the pooled data set used in the positron emission tomography (PET) study. Finally, we demonstrated an overlap of activations in the prefrontal region, previously reported as showing subtle across-language differences in activations (Kim et al., 1997). This was an area not revealed in Perani’s study.

How do we explain the results of studies showing differences in activations associated with native (L1) and second (L2) language processing? One possibility is that these differences reflect unbalanced processing load between languages when presented in sentence form. Existing studies reveal that cortical activation in single word tasks varies depending on word frequency (Büchel et al., 1998) and exposure duration (Price et al., 1996). With sentence studies, noun frequency (Keller et al., 1998, Neuroimage, abstract) and syntactic difficulty (Just et al., 1996; Keller et al., 1998, Neuroimage, abstract) produce variations in the spatial extent of activation in the temporal regions.

Another explanation is that fluent bilinguals differ from nonfluent bilinguals. This is suggested in the contrasting results obtained from a pair of activation studies in which auditory sentence stimuli were used (Dehaene et al., 1997; Perani et al., 1998). Intra- and interhemispheric
differences in language-related activations seen in the former study were eliminated in the latter study, which recruited fluent bilinguals.

Less fluent bilinguals utilize a word association strategy to gain access to concepts. In contrast, fluent bilinguals have been shown to access concepts directly from L2 without having to perform an internal translation through L1 (Kroll and deGroot, 1997). Both task demands (Thompson-Schill et al., 1997) and strategy (Reichle et al., 1998, Neuroimage, abstract) can influence the topography (as distinct from spatial extent) of cortical activation. We expect that in fluent compared to nonfluent bilinguals, a difference in access-to-concept mechanics will be reflected in dissimilar patterns of cortical activation. In a motor sequence learning experiment, the prefrontal and parietal cortices were more active during learning of new sequences, whereas the SMA was more active during performance of previously learned sequences (Jenkins et al., 1994). There is presently no direct imaging evidence for a parallel situation in language learning, although at the single word level, in pooled data and in already competent bilinguals, synonym generation and translation reportedly give rise to overlapping activations in L1 and L2 (Klein et al., 1995). The possibility of alterations in brain topography of activation with learning is suggested by changes in late components of ERP that occur as fluency is attained while mastering a miniature artificial language (McCandliss et al., 1997).

Chinese–English bilinguals have been shown to recognize code-switched words as quickly as monolingual English speakers (Li, 1996). Bilingual Stroop experiments show that picture naming in L1 is interfered with by the superimposition of an L2 distracter word on the picture (Smith and Kirsner, 1991; Smith, 1997). These observations suggest that fluency is accompanied by a convergence of conceptual representation. Our findings provide functional-anatomical support for the notion that common conceptual access is matched by the overlap of neuronal networks for comprehending sentences L1 and L2.

What prevents L1 and L2 from interfering with one another? Penfield (Penfield and Roberts, 1959) postulated an “automatic switch that allows the individual to turn from one language to another” (see also Green, 1986; Paradis, 1998). This idea can explain clinical phenomena such as the condition of alternate antagonistic cortical activation. We expect that in fluent compared to nonfluent bilinguals, a difference in access-to-concept mechanics will be reflected in dissimilar patterns of cortical activation. In a motor sequence learning experiment, the prefrontal and parietal cortices were more active during learning of new sequences, whereas the SMA was more active during performance of previously learned sequences (Jenkins et al., 1994). There is presently no direct imaging evidence for a parallel situation in language learning, although at the single word level, in pooled data and in already competent bilinguals, synonym generation and translation reportedly give rise to overlapping activations in L1 and L2 (Klein et al., 1995). The possibility of alterations in brain topography of activation with learning is suggested by changes in late components of ERP that occur as fluency is attained while mastering a miniature artificial language (McCandliss et al., 1997).

Chinese–English bilinguals have been shown to recognize code-switched words as quickly as monolingual English speakers (Li, 1996). Bilingual Stroop experiments show that picture naming in L1 is interfered with by the superimposition of an L2 distracter word on the picture (Smith and Kirsner, 1991; Smith, 1997). These observations suggest that fluency is accompanied by a convergence of conceptual representation. Our findings provide functional-anatomical support for the notion that common conceptual access is matched by the overlap of neuronal networks for comprehending sentences L1 and L2.

What prevents L1 and L2 from interfering with one another? Penfield (Penfield and Roberts, 1959) postulated an “automatic switch that allows the individual to turn from one language to another” (see also Green, 1986; Paradis, 1998). This idea can explain clinical phenomena such as the condition of alternate antagonistic recovery (Paradis et al., 1982). In this condition, the patient recovers only one language after a stroke. He then loses access to that language and the other language recovers. This cycle alternates, hence the name. Such a “switching model” has been criticized because it does not have much empirical support and does not adequately account for phenomena like faster recovery in related compared to unrelated languages (Sasanuma and Park, 1995). However, in the light of our findings, this model warrants reevaluation. Putative candidates for the hypothetical switch lie in the left prefrontal region. These are involved in selecting one verbal item from many (BA 9, 10) (Desmond et al., 1998) and in selecting one concept from competing ones (BA 9, 44, 45) (Thompson-Schill et al., 1997).

Imaging studies in American Sign Language (ASL) show right in addition to left hemisphere activations in subjects who learned ASL early but not later in life. These findings imply that cerebral organization for language may be determined by the structure and processing requirements of the language and that these are modulated by age of acquisition (Bavelier et al., 1998a, 1998b). We selected subjects exposed to Mandarin early in life.
in order to maximize the chances of uncovering right hemisphere utilization during Mandarin sentence processing. The overlapping activations observed in this study support the notion that it is the “spatialized syntax” (related to the use of dynamic hand movements rather than static written symbols) of ASL that contributes to right hemisphere activations. This suggests that additional cortical areas in either hemisphere may not be required to accommodate the structural differences that underlie different written languages.

More work is required to determine if our findings apply to sentence comprehension in late onset bilinguals and whether or not overlapping language representation is the end result or a prerequisite for fluency. For now, we can reasonably conclude that fluent bilinguals who have acquired both languages in early childhood utilize common neuroanatomical areas while comprehending written sentences, irrespective of the surface characteristics inherent in each language.

Experimental Procedures

Behavioral Pilot and Experiment 1
A pilot behavioral study was conducted on 15 volunteers (19-26 years of age) who gave informed consent. fMRI Experiment 1 utilized the same experimental paradigm. Subjects were fluent in written English and Mandarin. They were all Singaporean undergraduates of Chinese descent who were exposed to both English and Mandarin prior to the age of 6. All subjects scored at least a B grade in middle school examinations for both languages and used both languages in daily life. A secondary measure of equivalency of written language competence was that all subjects scored 70% in accuracy or better in the experimental task in both languages.

Mandarin and English words were presented on a monitor using Chinese Language Kit (Apple Computer, Cupertino, CA) and MacStim 2.2.7 (D. Darby, Melbourne, Australia). The stimuli consisted of 60 sentences in English and 60 in Mandarin. Mandarin sentences were translated from the English ones. Every subject was instructed to read and understand each sentence and then respond to the ensuing probe question by pressing one button for “true” and another button for “false” with the right hand.

Sentences were presented in a similar syntactic form for each language and were modeled after those used by Just and Carpenter (Just et al., 1996). In English, sentences were “subject object relative”; e.g., “The speech that the minister gave angered the reporter.” In Mandarin, the “subject object relative” sentence was “bu zhang suo fa biao de yan lun chu nu le ji zhe” (hanyu pinyin equivalent of the Mandarin statement). The subjects and objects (“agents” and “themes” in some conventions) of each sentence were chosen so that the sentence could be reversed and permutations of the probe question were plausible. In the example, four probe questions were possible:

- The minister gave the speech? (true)
- The speech angered the reporter? (true)
- The reporter gave the speech? (false)
- The speech angered the minister? (false)
The minister angered the reporter? (false)

For Mandarin, a similar organization of sentence and probe was used. Probe questions were based on the first clause (the minister gave the speech) or the second clause (the speech angered the reporter) and were counterbalanced. The number of true and false responses was also counterbalanced. Sentences were randomized within 30 s blocks comprising five sentence-probe pairs. Each sentence was presented for 3 s followed by a 0.5 s fixation interval and each probe for 2 s followed by a 0.5 s fixation interval (Figure 1). The short timing was intended to reduce the likelihood of translation during task performance. Four blocks of sentences were presented in each of six trials ("runs"). Blocks of English sentences and Mandarin sentences were alternated and the order of presentation was counterbalanced across runs. Across languages, semantically matched sentences were used. As such, the second half of each experiment used stimuli that were translated versions of those presented previously. There was a two-run separation between cross-language repetition of stimuli.

Thirteen right-handed subjects (20-26 years of age) who gave written consent participated in Experiment 1. The language backgrounds and fluency in both languages of these subjects were similar to that of the behavioral pilot study. Of the original thirteen, three subjects were rejected as a result of failure to meet the accuracy of response criteria. One subject had exclusively right hemisphere activations (which overlapped across languages as well). We omitted this subject as we wish to discuss her in a separate communication. Nine subjects were finally analyzed. Sentence stimuli were presented through a fiber-optic projector system (Avotec, Jensen Beach, FL). The stimuli used in Experiment 1 were identical to those used in the behavioral study (Figure 1, top).

A two-button mouse was used to collect behavioral data concurrent with imaging. The physical set up of the response system was the cause for the slight delay in responses collected from the scanner compared to the behavioral study. This problem was corrected in Experiment 2.

Experiment 2

Five subjects (18-22 years of age) participated in this experiment, two of whom had previously participated in Experiment 1. The time interval between the second and first experiments was more than a month in each of the repeated subjects. The stimuli used were identical with the previous experiment except for three modifications (Figure 1, bottom). These modifications sought to reduce contributions from early visual processing and to counter concern that substantial semantic processing could occur during "neutral" tasks like fixation (Binder et al., 1999). A control task involving a string of nonsense Tamil characters and a "probe question" comprising nonsense Tamil characters was implemented. Subjects were asked to scan the Tamil-like pseudo-characters as if reading them and to respond with alternating left and right mouse button presses to successive "probe questions." Care was taken to ensure that all scripts subtended the same visual angle in the vertical direction. The control task appeared for 40 s. Blocks containing sentence stimuli were increased in duration from 30 to 60 s. In contrast to Experiment 1, probes on the first and second clause of each sentence were randomized within the same block.

Apparatus and Scanning Procedure

Scanning was performed in a 2.0T Bruker Tomikon S200 system (Bruker, Ettlingen, Germany) fitted with a 30 mT/m gradient system. A gradient-echo echoplanar imaging (EPI) sequence with the following parameters was used: effective TE 40 ms, TR 2000 ms, FOV 23 x 23 cm, pixel matrix 128 x 64 (128 pixels in the A-P direction), and 17 contiguous oblique axial slices approximately parallel to the AC-PC line 4 mm thick (with a 2 mm gap). Six hundred images were collected at each of the slice locations, in six separate imaging runs. High-resolution, T1-weighted anatomical reference images were obtained as a set of 128 contiguous axial slices with a reformatted matrix of 256 x 256 pixels, using a 3D spoiled gradient-echo sequence. Head motion was reduced using a bite bar system (Biomat, Singapore).

Image Processing and Data Analysis

An inherent difficulty of EPI imaging sequences is that a phase difference and relative k-space shift of even and odd gradient echoes arises from the time reversal of alternate k-space lines due to opposite polarities of the readout gradients. This results in the appearance of ghost images, shifted by one half of the field of view in the phase-encoding direction. As the ghosts interfere with the original image and cause loss of amplitude, a phase correction was applied to suppress them (Buonocore and Gao, 1997; Hennel, 1998). In this technique, the reconstructed EPI image was decomposed into a sum of two complex images, one reconstructed from the odd lines of k-space and another from the even lines. Phase errors were extracted from pixels in the nonoverlapping regions of the parent image and used for reduction of the ghost artifacts.

Images were subsequently processed using MedX 2.11/3.0 (Sensor Systems, Bethesda, MD). Rigid body motion correction was performed using automated image registration (AIR) (Woods et al., 1992). To reduce the effect of variation of MR signal between runs, the resulting motion-corrected images were globally normalized to an empirically determined median value of 6500 units.

To facilitate coregistration of the EPI images with the high-resolution anatomical image, removal of the skull and dura were performed with a semi-automated segmentation utility. The resultant images were transformed into Talairach space (Talairach and Tournoux, 1988) using a linear scaling method, and the transformation matrix was saved. The first time point of the normalized, motion-corrected images was then registered to the high-resolution, skull-stripped anatomical image using AIR. The transformation matrix for this procedure was again saved.

Functional images were separated into English, Mandarin, and fixation groups. Images were shifted 4 s to allow for a time lag in the rise of blood oxygen level-dependent (BOLD) signal following presentation of word stimuli. Unpaired t tests were then applied to the groups of images. Three comparisons using a boxcar function to model subject responses were made: English sentences versus fixation, Mandarin versus fixation, and English versus Mandarin. We cross-checked the activation maps derived from pooling fixation epochs by comparing these with maps separately derived from comparing task-related signals against epochs of fixation preceding and following each task. A Gaussianized t test was used to compute Z maps, and Z score thresholds of 3.5 and 5 were used to generate activation maps. Activated voxels in regions of interest were checked to see if there was an appropriate time-locked variation in MR signal in relation to stimulus presentation. The somewhat conservative Z score threshold of 5 was initially used in order to minimize false-positive activations. Concern that this would obscure subtle differences in across-language activations persuaded us to use a lower threshold of 3.5. Z score maps were transformed into Talairach space and coregistered with the high-resolution anatomical images. Averaging Talairach-transformed statistical and structural images from each subject separately and then recomputing the statistical map created pooled activation maps.

A peak detection routine within MedX was used to determine activation maxima. Peaks within the frontal and temporal areas on the left hemisphere were tabulated to give an idea of the extent of coactivation of brain areas in the two languages. Voxels exceeding a statistical threshold of Z = 3.5 and 5 in the lateral frontal regions were counted. This included frontal activations in Brodmann's areas 47, 44, 45, 9, and 6. An AI (Binder et al., 1996) of activated voxels in the frontal region was calculated for each of the repeated subjects. The stimuli used were in MR signal in relation to stimulus presentation. The somewhat subtle differences in across-language activations persuaded us to minimize false-positive activations. Concern that this would obscure subtle differences in across-language activations persuaded us to use a lower threshold of 3.5. Z score maps were transformed into Talairach space and coregistered with the high-resolution anatomical images. Averaging Talairach-transformed statistical and structural images from each subject separately and then recomputing the statistical map created pooled activation maps.

A peak detection routine within MedX was used to determine activation maxima. Peaks within the frontal and temporal areas on the left hemisphere were tabulated to give an idea of the extent of coactivation of brain areas in the two languages. Voxels exceeding a statistical threshold of Z = 3.5 and 5 in the lateral frontal regions were counted. This included frontal activations in Brodmann's areas 47, 44, 45, 9, and 6. An AI (Binder et al., 1996) of activated voxels in the frontal region was calculated for each of the repeated subjects. The stimuli used were in MR signal in relation to stimulus presentation. The somewhat subtle differences in across-language activations persuaded us to minimize false-positive activations. Concern that this would obscure subtle differences in across-language activations persuaded us to use a lower threshold of 3.5. Z score maps were transformed into Talairach space and coregistered with the high-resolution anatomical images. Averaging Talairach-transformed statistical and structural images from each subject separately and then recomputing the statistical map created pooled activation maps.

A peak detection routine within MedX was used to determine activation maxima. Peaks within the frontal and temporal areas on the left hemisphere were tabulated to give an idea of the extent of coactivation of brain areas in the two languages. Voxels exceeding a statistical threshold of Z = 3.5 and 5 in the lateral frontal regions were counted. This included frontal activations in Brodmann's areas 47, 44, 45, 9, and 6. An AI (Binder et al., 1996) of activated voxels in the frontal region was calculated for each of the repeated subjects. The stimuli used were in MR signal in relation to stimulus presentation. The somewhat subtle differences in across-language activations persuaded us to minimize false-positive activations. Concern that this would obscure subtle differences in across-language activations persuaded us to use a lower threshold of 3.5. Z score maps were transformed into Talairach space and coregistered with the high-resolution anatomical images. Averaging Talairach-transformed statistical and structural images from each subject separately and then recomputing the statistical map created pooled activation maps.
References


