

# 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

||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 Opler, 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

# To whom correspondence should be addressed (e-mail: mchee@pacific.net.sg).

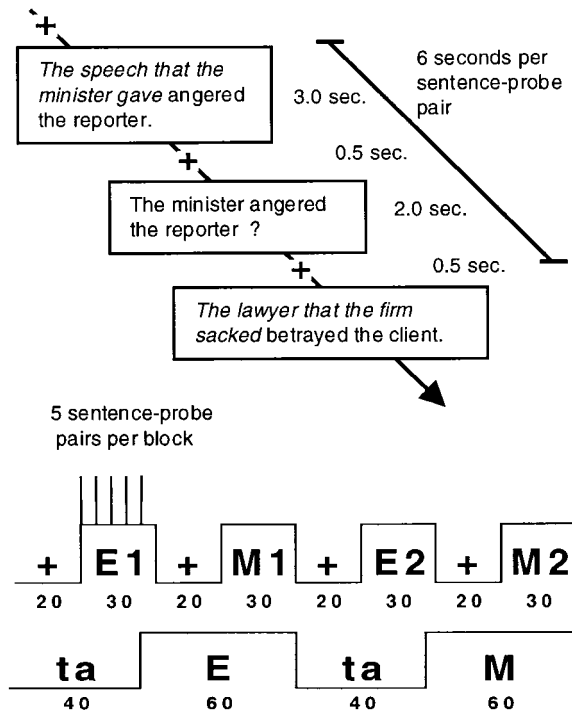


Figure 1. Experimental Design

The design and block timing of both fMRI experiments are shown. "E1" refers to English sentence-probe pairs testing the first clause (in italics), and "E2" refers to English sentence-probe pairs testing the second clause. "E" and "M" denote blocks in which probes on both clauses are randomized within the same block. Block duration (in seconds) appears below each block. "Ta" denotes Tamil-like pseudo-word stimuli. Block duration (in seconds) appears below each frame. The probe question in this example tests the second clause. A similar arrangement of stimuli was used for Mandarin sentences.

uncovered could be attributed to surface differences between the languages.

## Results

### Behavioral and fMRI Experiments

Experiments in which subjects evaluated sentences with a fixed structure were conducted. Subjects responded to a probe question that followed each sentence by indicating "true" or "false" with a two-button mouse. Prior to two sets of fMRI experiments, a pilot set of purely behavioral experiments was conducted to ensure that the sentence task was of comparable difficulty in both languages. In the first set of fMRI experiments, sentence-probe pairs (Figure 1) were presented at fixed intervals within alternating English and Mandarin blocks. The question either probed the first or the second clause of the presented sentence. Fixation served as the baseline condition. The comparison of sentence task and fixation was intended to engage the entire set of processes related to sentence processing. In the second set of fMRI experiments, the duration of presentation for the stimulus and control blocks was lengthened. Probe questions tested either clause. A new control task using Tamil-like pseudo-characters was implemented. This was expected to markedly reduce activations related to early visual processing of word-like stimuli as

Table 1. Accuracy and Reaction Time Data

	Accuracy (% Correct)		Reaction Time (ms)	
	Mean	SD	Mean	SD
Behavioral pilot (n = 15)				
English	88.4	5.0	1272	217
Mandarin	89.2	5.5	1271	160
fMRI Experiment 1 (n = 9)				
English	89.6	5.5	1485	126
Mandarin	88.4	6.3	1502	148
fMRI Experiment 2 (n = 5)				
English	89.7	10.5	1236	140
Mandarin	87.0	10.4	1354	175

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  $F_s < 1$ ).

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

Table 2. Talairach Coordinates of Left Frontal and Temporal Activations in Individuals from Experiment 1

Subject	Mandarin			English			Gyrus	Brodmann's Area		
	Z score	x	y	z	Z score	x			y	z
GR	24.8	-44	10	36	11.4	-46	10	36	L MFG	9
	16.1	-40	14	28	10.0	-40	14	30	L IFG	9, 44
KC	29.9	-40	8	48	30.0	-40	8	48	L MFG	8
	28.9	-42	18	34	28.8	-42	18	34	L MFG	9
	26.5	-52	12	14	22.3	-52	12	14	L IFG	44
LI					27.6	-46	16	44	L MFG	8
	25.4	-46	20	30	25.2	-46	20	30	L MFG	9
					11.5	-46	-42	20	L posterior STG	22
SH	19.6	-52	-24	14	11.2	-54	-22	14	L mid STG	22
	9.1	-44	12	18	8.7	-44	12	18	L IFG	44
	8.8	-38	10	24	8.2	-38	10	24	L IFG	44
	8.0	-50	-24	-2	8.2	-50	-24	-2	L STG	21/22
TJ	7.8	-42	-36	0	7.7	-42	-36	0	L STG	21/22
	26.4	-40	6	28	18.0	-40	14	14	L IFG	44
	16.4	-40	4	40	17.0	-40	4	40	L MFG	6
	11.7	-40	14	10	18.0	-40	14	14	L IFG	44
	11.4	-56	-26	-4	11.5	-56	-26	-4	L MTG	21
YE	10.1	-50	-50	12	10.0	-50	-50	12	L posterior STG	22
	30.8	-42	18	26	31.9	-42	18	26	L IFG, MFG	44, 9
	29.1	-46	12	44	29.1	-46	12	44	L MFG	8
	10.5	-42	6	26	10.8	-42	6	26	L IFG	44
	10.2	-54	-42	10	10.7	-54	-42	10	L posterior STG	22
YU	27.5	-36	10	38	28.4	-36	10	38	L IFG	44, 9
	26.6	-34	18	30	25.2	-32	18	30	L MFG	9
	10.1	-42	-38	14	8.2	-40	-38	14	L posterior STG	22
ZI	9.6	-42	22	30	9.7	-40	20	30	L MFG	9
	9.1	-44	18	42	9.5	-44	18	42	L MFG	9
	7.9	42	12	28					L IFG	9, 44
YO	22.4	-40	2	32	11.3	-40	2	32	L MFG	9, 45
	20.6	-42	20	30	11.6	-42	18	30	L MFG	9, 45
	10.1	-54	-42	8	7.7	-54	-40	8	L posterior STG	22

IFG, inferior frontal gyrus; MFG, middle frontal gyrus; STG, superior temporal gyrus; MTG, middle temporal gyrus.

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 pseudo-word 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).

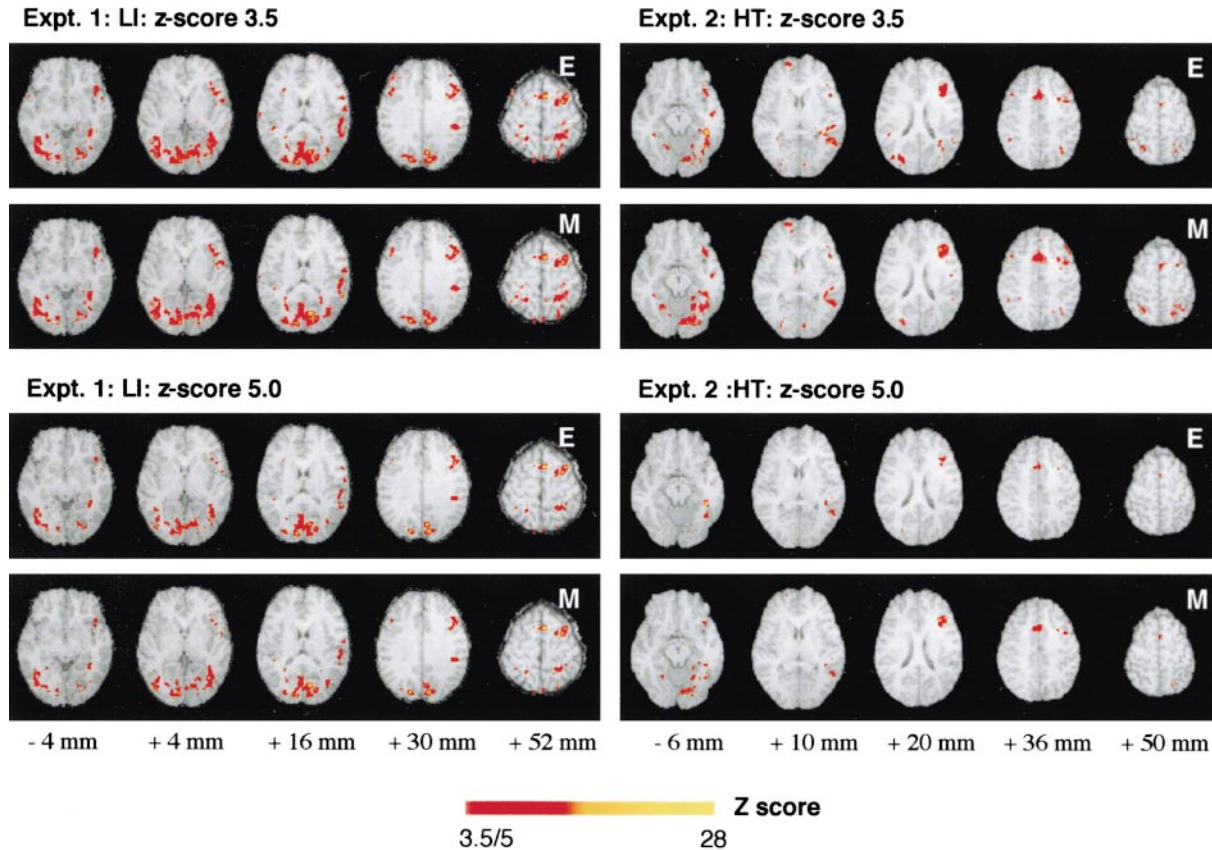


Figure 2. Activation Maps from Representative Individuals from Experiments 1 and 2 Shown at Two Different Statistical Thresholds  
The subject's left hemisphere is on the right side of each image. Areas activated above a Z score of 3.5 and 5.0 are shown in two subjects, LI and HT. "E" and "M" denote maps relating to English and Mandarin, respectively. The figures below each set of strips indicate the distance above or below the intercommisural plane.

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

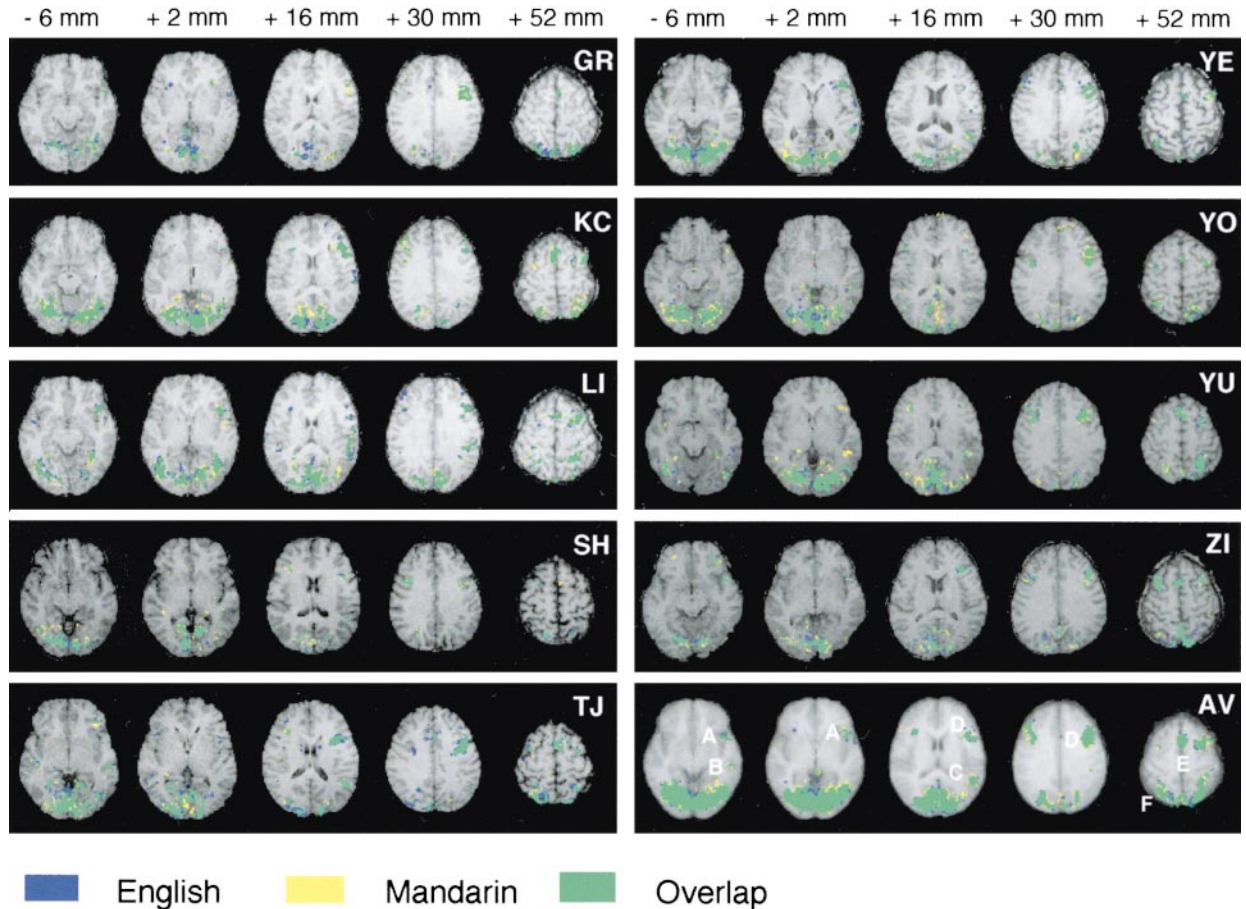


Figure 3. Individual and Pooled Activation Maps from Experiment One

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

	Frontal Activations (voxels)				Temporal Activations (voxels)				Asymmetry Index	
	Mandarin		English		Mandarin		English		Mandarin	English
	R	L	R	L	R	L	R	L		
Experiment 1										
GR	5	488	21	422	0	0	0	0	0.98	0.91
KC	103	570	87	469	0	0	0	0	0.69	0.69
LI	152	691	197	722	0	347	11	387	0.64	0.57
SH	92	156	43	120	0	23	0	22	0.26	0.47
TJ	298	863	324	1015	23	105	23	96	0.49	0.52
YE	36	731	51	827	0	121	0	108	0.91	0.88
YU	147	710	132	293	1	370	0	29	0.66	0.38
ZI	70	181	107	210	0	0	0	0	0.44	0.32
YO	179	520	179	406	0	30	0	7	0.49	0.39
AV3.5									0.48	0.46
AV5.0									0.73	0.67
Experiment 2										
SH	20	42	0	30	70	182	31	94	0.35	1.00
PL	6	222	3	324	2	46	2	26	0.95	0.98
HT	27	267	3	90	5	77	0	63	0.82	0.94
LI	0	28	0	41	2	81	4	171	1.00	1.00
DE	46	380	15	441	0	54	0	86	0.78	0.93
AV3.5									0.73	0.88
AV5.0									0.97	0.99

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 = sum [voxels (L - R)]/sum [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 (Buchel 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

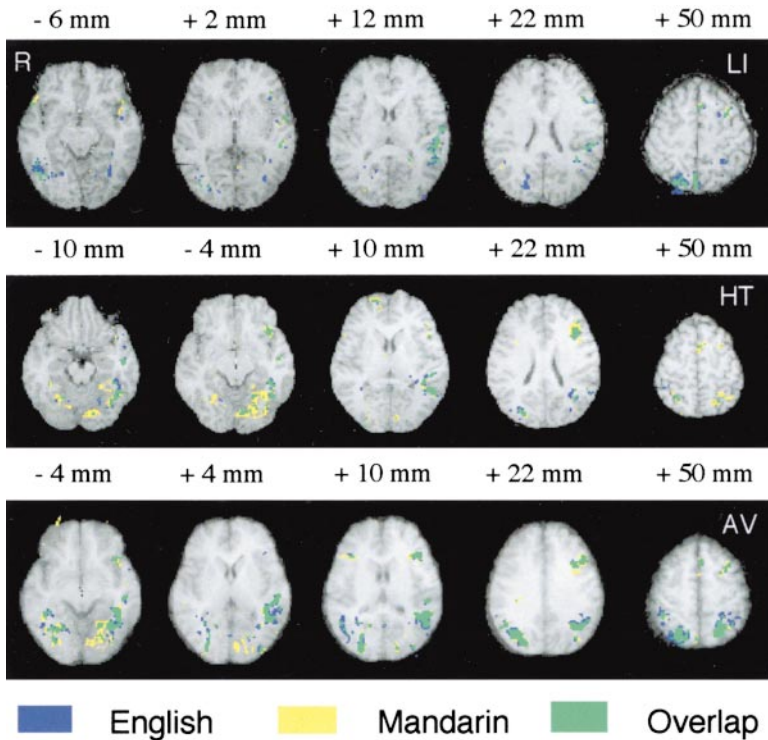


Figure 4. Individual and Pooled Activation Maps from Experiment Two

Montage of activations of two representative subjects participating in Experiment 2 (where Tamil-like pseudo-words acted as 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. Distances above or below the intercommissural plane are indicated. Subjects' initials appear on the upper right hand corner of each strip. "AV" represents the pooled data set.

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 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

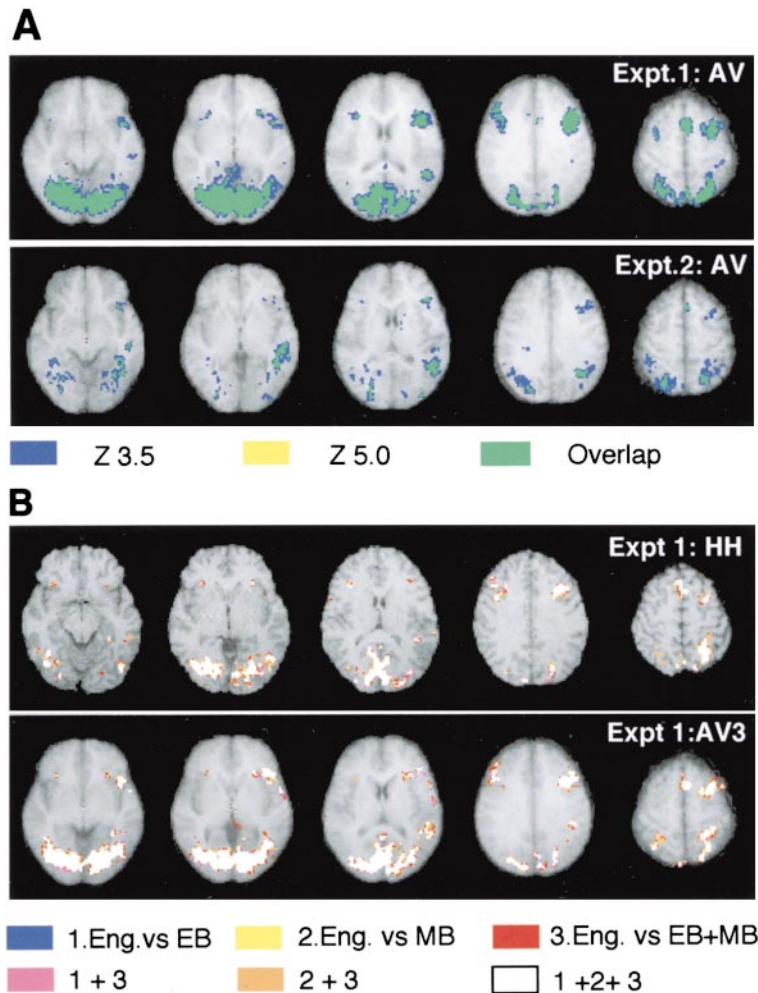


Figure 5. The Effects of Statistical Threshold and Choice of Baseline Epoch on Activation (A) Pooled activations associated with English sentence processing at two different thresholds (Z scores, 3.5 and 5.0). Averaged data from Experiment 1 appear in the upper panel and data from Experiment 2 in the lower panel.

(B) The effect of switching the baseline for a representative subject from Experiment 1 (top) and the average of three subjects' data (YE, HH, LI; bottom) are shown. Activations associated with English sentence comprehension were compared to the period of fixation following the English task (blue), the Mandarin task (yellow), and a combined baseline (red). The white areas indicate complete overlap of activations irrespective of the source of the baseline used.

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 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 can 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 × 23 cm, pixel matrix 128 × 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 × 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 language. Correlation was computed between English and Mandarin AIs for Experiment 1.

## Acknowledgments

This work was supported by the National Medical Research Council of Singapore, Grant NMRC 98/00270. D. C. received support from NINCD Grant DC02146. Xue Jie provided helpful suggestions. Chung Sew Meng of BIOMAT, Singapore, designed the head restraint system. Robert Savoy assisted in the design of the subject response unit. Jia Jia Hoon and Marilyn Chee helped run experiments and prepare the manuscript. We acknowledge the editor and anonymous reviewers for helpful suggestions and comments.

Received November 30, 1998; revised April 20, 1999.

## References

- Albert, M., and Obler, L. (1978). *The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism* (New York: Academic Press).
- 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.
- Bavelier, D., Corina, D., Jezzard, P., Clark, V., Karni, A., Lalwani, A., Rauschecker, J., Braun, A., Turner, R., and Neville, H. (1998a). Hemispheric specialization for English and ASL: left invariance–right variability. *Neuroreport* *9*, 1537–1542.
- Bavelier, D., Corina, D.P., and Neville, H.J. (1998b). Brain and language: a perspective from sign language. *Neuron* *21*, 275–278.
- Binder, J.R., Swanson, S.J., Hammeke, T.A., Morris, G.L., Mueller, W.M., Fischer, M., Benbadis, S., Frost, J.A., Rao, S.M., and Houghton, V.M. (1996). Determination of language dominance using functional MRI: a comparison with the WADA test. *Neurology* *46*, 978–984.
- Binder, J., Frost, J., Hammeke, T., Bellgowan, P., Rao, S., and Cox, R. (1999). Conceptual processing during the conscious resting state: a functional MRI study. *J. Cogn. Neurosci.* *11*, 80–95.
- Büchel, C., Holmes, A.P., Rees, G., and Friston, K.J. (1998). Characterizing stimulus–response functions using nonlinear regressors in parametric fMRI experiments. *Neuroimage* *8*, 140–148.
- Buonocore, M.H., and Gao, L. (1997). Ghost artifact reduction for echo planar imaging using image phase correction. *Magn. Reson. Med.* *38*, 89–100.
- Caplan, D., Hildebrandt, N., and Makris, N. (1996). Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain* *119*, 933–949.
- Chee, M.W.L., Tan, E., and Thiel, T. (1999). Mandarin and English single word processing studied with fMRI. *J. Neurosci.* *19*, 3050–3056.
- 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.
- Desmond, J.E., Gabrieli, J.D., and Glover, G.H. (1998). Dissociation of frontal and cerebellar activity in a cognitive task: evidence for a distinction between selection and search. *Neuroimage* *7*, 368–376.
- Fiez, J.A., Raife, E.A., Balota, D.A., Schwarz, J.P., Raichle, M.E., and Petersen, S.E. (1996). A positron emission tomography study of the short-term maintenance of verbal information. *J. Neurosci.* *16*, 808–822.
- Gabrieli, J.D., Poldrack, R.A., and Desmond, J.E. (1998). The role of left prefrontal cortex in language and memory. *Proc. Natl. Acad. Sci. USA* *95*, 906–913.
- Green, D. (1986). Control, activation, and resource: a framework and model. *Brain Lang.* *27*, 210–233.
- Hennel, F. (1998). Image-based reduction of artifacts in multishot echo-planar imaging. *J. Magn. Reson.* *134*, 206–213.
- Jenkins, I.H., Brooks, D.J., Nixon, P.D., Frackowiak, R.S.J., and Passingham, R.E. (1994). Motor sequence learning: a study with positron emission tomography. *J. Neurosci.* *14*, 3775–3790.
- Just, M., Carpenter, P., Keller, T., Eddy, W., and Thulborn, K. (1996). Brain activation modulated by sentence comprehension. *Science* *274*, 114–116.
- Kelley, W.M., Miezin, F.M., McDermott, K.B., Buckner, R.L., Raichle, M.E., Cohen, N.J., Ollinger, J.M., Akbudak, E., Conturo, T.E., Snyder, A.Z., and Petersen, S.E. (1998). Hemispheric specialization in human dorsal frontal cortex and medial temporal lobe for verbal and nonverbal memory encoding. *Neuron* *20*, 927–936.
- 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., Meyer, E., and Evans, A. (1995). The neural substrates underlying word generation: a bilingual functional-imaging study. *Proc. Natl. Acad. Sci. USA* *92*, 2899–2903.
- Klein, D., Zatorre, R., Milner, B., Meyer, E., and Evans, A. (1994). Left putaminal activation when speaking a second language: evidence from PET. *Neuroreport* *5*, 2295–2297.
- Koyama, S., Kakigi, R., Hoshiyama, M., and Kitamura, Y. (1998). Reading of Japanese Kanji (morphograms) and Kana (syllabograms): a magnetoencephalographic study. *Neuropsychologia* *36*, 83–98.
- 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: Psycholinguistic perspectives*, J.F. Kroll and A.M.B. deGroot, eds. (Mahwah, NJ: Lawrence Erlbaum Associates), pp. 169–199.
- Li, P. (1996). Spoken word recognition of code-switched words by Chinese–English bilinguals. *J. Mem. Lang.* *35*, 757–774.
- Mazoyer, B., Dehaene, S., Tzourio, N., Frak, V., Murayama, N., Cohen, L., Levrier, O., Salamon, G., Syrota, A., and Mehler, J. (1993). The cortical representation of speech. *J. Cogn. Neurosci.* *5*, 467–479.
- McCandliss, B.D., Posner, M.I., and Givon, T. (1997). Brain plasticity in learning visual words. *Cogn. Psychol.* *33*, 88–110.
- McClelland, J.L., St. John, M., and Taraban, R. (1989). Sentence comprehension: a parallel distributed processing approach. *Lang. Cogn. Process.* *4*, 287–335.
- Müller, R.A., Rothermel, R.D., Behen, M.E., Muzik, O., Mangner, T.J., and Chugani, H.T. (1997). Receptive and expressive language activations for sentences: a PET study. *Neuroreport* *8*, 3767–3770.
- Neville, H.J., and Bavelier, D. (1998). Neural organization and plasticity of language. *Curr. Opin. Neurobiol.* *8*, 254–258.
- Nobre, A., Allison, T., and McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature* *372*, 260–263.
- Ojemann, G., and Whitaker, H. (1978). The bilingual brain. *Arch. Neurol.* *35*, 409–412.
- Osterhout, L., McLaughlin, J., and Bersick, M. (1997). Event-related brain potentials and human language. *Trends Cogn. Sci.* *1*, 203–209.
- Paradis, M. (1995). *Aspects of Bilingual Aphasia* (Oxford: Pergamon Press).
- Paradis, M. (1998). Language and communication in multilinguals. In *Handbook of Neurolinguistics*, B. Stemmer and H. Whitaker, eds. (San Diego, CA: Academic Press), pp. 417–430.
- Paradis, M., Goldblum, M., and Abidi, R. (1982). Alternate antagonism with paradoxical translation behavior in two bilingual aphasic patients. *Brain Lang.* *15*, 55–69.
- Penfield, W., and Roberts, L. (1959). *Speech and Brain Mechanisms* (Princeton, NJ: Princeton University Press), pp. 253.
- 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., Betti-nardi, 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.
- Poeppl, D. (1996). A critical review of PET studies of phonological processing. *Brain Lang.* *55*, 317–351.
- Price, C.J., Moore, C.J., and Frackowiak, R.S. (1996). The effect of varying stimulus rate and duration on brain activity during reading. *Neuroimage* *3*, 40–52.
- Rapport, R., Tan, C., and Whitaker, H. (1983). Language function and dysfunction among Chinese- and English-speaking polyglots: cortical stimulation, Wada testing, and clinical studies. *Brain Lang.* *18*, 342–366.
- Sasanuma, S., and Park, H.S. (1995). Patterns of language deficits in two Korean–Japanese bilingual aphasic patients—a clinical report. In *Aspects of Bilingual Aphasia*, M. Paradis, ed. (Oxford: Pergamon Press), pp. 111–122.
- Smith, M.C. (1997). How do bilinguals access lexical information? In *Tutorials in Bilingualism: Psycholinguistic Perspectives*, J.F. Kroll and A.M.B. deGroot, eds. (Mahwah, NJ: Lawrence Erlbaum Associates), pp. 145–167.
- Smith, M.C., and Kirsner, K. (1991). Language and orthography as irrelevant features in colour–word and picture–word Stroop interference. *Q. J. Exp. Psychol. A* *34*, 153–170.

- Smith, E.E., Jonides, J., Marshuetz, C., and Koeppel, R.A. (1998). Components of verbal working memory: evidence from neuroimaging. *Proc. Natl. Acad. Sci. USA* *95*, 876–882.
- Stromswold, K., Caplan, D., Alpert, N., and Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* *52*, 452–473.
- Sugishita, M., Otomo, K., Kabe, S., and Yunoki, K. (1992). A critical appraisal of neuropsychological correlates of Japanese ideogram (Kanji) and phonogram (Kana) reading. *Brain* *115*, 1563–1585.
- Talairach, J., and Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain* (New York: Thieme Medical Publishers).
- 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.
- Weber-Fox, C., and Neville, H. (1997). Maturation constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *J. Cogn. Neurosci.* *8*, 231–256.
- Woods, R.P., Cherry, S.R., and Mazziotta, J.C. (1992). Rapid automated algorithm for aligning and reslicing PET images. *J. Comput. Assist. Tomogr.* *16*, 620–633.