

# Modality Independence of Word Comprehension

James R. Booth,<sup>1–3\*</sup> Douglas D. Burman,<sup>1</sup> Joel R. Meyer,<sup>2</sup>  
Darren R. Gitelman,<sup>3,4</sup> Todd B. Parrish,<sup>3,5</sup> and M. Marsel Mesulam<sup>3,4</sup>

<sup>1</sup>Department of Communication Sciences and Disorders, Northwestern University, Evanston, Illinois

<sup>2</sup>Department of Radiology, Evanston Hospital, Evanston, Illinois

<sup>3</sup>Cognitive Brain Mapping Group, Northwestern University, Chicago, Illinois

<sup>4</sup>Department of Neurology, Northwestern University Medical School, Chicago, Illinois

<sup>5</sup>Department of Radiology, Northwestern University Medical School, Chicago, Illinois

**Abstract:** Functional magnetic resonance imaging (fMRI) was used to examine the functional anatomy of word comprehension in the auditory and visual modalities of presentation. We asked our subjects to determine if word pairs were semantically associated (e.g., table, chair) and compared this to a reference task where they were asked to judge whether word pairs rhymed (e.g., bank, tank). This comparison showed task-specific and modality-independent activation for semantic processing in the heteromodal cortices of the left inferior frontal gyrus (BA 46, 47) and left middle temporal gyrus (BA 21). There were also modality-specific activations in the fusiform gyrus (BA 37) for written words and in the superior temporal gyrus (BA 22) for spoken words. Our findings are consistent with the hypothesis that word form recognition (lexical encoding) occurs in unimodal cortices and that heteromodal brain regions in the anterior as well as posterior components of the language network subserve word comprehension (semantic decoding). *Hum. Brain Mapping* 16:251–261, 2002. © 2002 Wiley-Liss, Inc.

**Key words:** semantic; meaning; spoken; written; auditory; visual; fMRI; lexical

## INTRODUCTION

The comprehension of language requires the conversion of sensory input patterns into word forms and the subsequent linkage of these word forms to the distributed associations that encode their meaning [Mesulam, 1998]. Neuroimaging studies of semantic processing have generally confined themselves to a

single modality of input. Most of these studies have examined the comprehension of written words [Friederici et al., 2000; Kapur et al., 1996; Mummery et al., 1998; Poldrack et al., 1999; Price et al., 1997; Pugh et al., 1996; Wagner et al., 1998]. Fewer studies have examined the comprehension of spoken words [Binder et al., 1997; Demonet et al., 1992]. A review of activation patterns in these studies shows that the left inferior frontal gyrus and the left middle temporal gyrus appear to be activated by either input modality. A more definitive statement concerning modality independent participation of these areas in word comprehension, however, requires experiments where written and spoken words are presented to the same subject.

Two studies have compared the comprehension of written words to the comprehension of spoken words

\*Correspondence to: James R. Booth, Department of Communication Sciences and Disorders, Northwestern University, 2299 North Campus Drive, Evanston, Illinois, 60208-3560.

E-mail: j-booth@nwu.edu

Received for publication 11 December 2001; accepted 19 April 2002

DOI 10.1002/hbm.10054

in the same subjects. Petersen et al. [1988, 1989] asked participants to generate verbs in response to written or spoken words [Petersen et al., 1988, 1989]. The baseline for the visually presented word task was reading aloud and the baseline for the auditorily presented word task was the repetition of heard words. As compared to their respective baselines, the semantic processing of only visually presented words produced activation in the left inferior frontal gyrus. Semantic tasks in neither input modality produced activation in the left middle temporal gyrus. In another study, Chee et al. [1999] asked participants to determine whether written or spoken words were concrete or abstract [Chee et al., 1999]. The baseline for the spoken task was syllable counting of heard words and the baseline for the written task was an upper versus lower case judgment task of visually presented words. As compared to their respective baselines, the semantic processing of both written and spoken words activated the left inferior frontal gyrus, but only the semantic processing of spoken words activated the left middle temporal gyrus.

This brief review shows that multi-modality studies are somewhat at odds with single-modality studies and that there is currently no consensus on the modality independent substrate of word comprehension. Contradictory results across studies may have resulted from differences in task characteristics or baseline conditions. In our study, rhyming tasks were used as a baseline for semantic processing in the visual as well as auditory modalities. Our goal was to identify the modality-specific and modality-independent substrates of lexical encoding and semantic processing.

## MATERIALS AND METHODS

### Participants

Thirteen adults (median age = 24.6 years, range = 20–35 years) participated in this study. Eleven of the adults (10 women, 1 man) completed both the visual and auditory tasks; the other two (both men) completed tasks in only one modality. All participants were right-handed according to a 10-item Likert-scale questionnaire (median = 89, range = 65–100). All participants were undergraduate or graduate students at Northwestern University, and were given an informal interview to ensure that they did not have a history of intelligence, reading, or oral-language deficits. All participants were native English speakers and had normal hearing and normal or corrected-to-normal vision. All participants were free of neurological

diseases or psychiatric disorders and were not taking medication affecting the central nervous system. Research was conducted according to the Code of Ethics of the World Medical Association (Declaration of Helsinki). The Institutional Review Board at Northwestern University and Evanston Northwestern Healthcare Research Institute approved the informed consent procedures.

### Functional activation tasks

In all tasks, three stimuli were presented sequentially and the participant had to determine whether the final stimulus matched either of the two previous stimuli according to a predefined criterion. Sixty percent of the trials involved a match and 40% involved a non-match. Half of the matching trials involved a match to the first stimulus and half involved a match to the second stimulus. If there was a match, the participant pressed a button with the index finger; if there was no match, the participant pressed a different button with the middle finger.

Each task lasted 9 min and consisted of 10 blocks of 54 sec each including a 4-sec single-word instruction screen at the beginning of each block. Five word judgment blocks alternated with the five control blocks. Each word was presented for 800 msec followed by a 200 msec blank interval. A yellow fixation cross (+) appeared on the screen after the third stimulus, indicating the need to make a response during the subsequent 2,000 msec interval. Participants were told that they could respond before the yellow cross (+) appeared on the screen. Each trial lasted a total of 5,000 msec and there were 10 trials in each block.

#### Visual meaning task

Participants determined whether a final word was associated with one of two preceding words [Nelson et al., 1994]. Half of the related pairs had a high association (e.g., answer, question) and half had a low association (e.g., hate, like). There was no overlap in free association values for the high associates (median = 0.59) and the low associates (median = 0.27).

#### Visual rhyming task

Participants determined whether the final word rhymed with either of the first two words. Half of the target trials contained a target word that rhymed and was orthographically similar to one of the preceding two words (e.g., seat, heat). The other half contained a

**TABLE I. Means and standard errors for accuracy and reaction time in the word judgment and control tasks in the visual and auditory modality for the practice and fMRI sessions\***

	Practice				fMRI			
	Accuracy (%)		RT (ms)		Accuracy (%)		RT (ms)	
Visual								
Meaning	96.7	5.2	1028	119	96.5	5.3	958	95
Control	95.7	5.9	788	74	99.0	2.9	761	72
Rhyming	96.2	5.6	901	107	96.5	5.3	898	87
Control	95.8	5.8	807	83	97.0	4.9	744	70
Auditory								
Meaning	98.0	4.1	1099	112	97.7	4.4	1105	110
Control	93.1	7.3	996	97	94.8	6.4	940	79
Rhyming	98.0	4.1	979	105	97.5	4.5	956	75.5
Control	92.4	7.6	1012	101	95.8	5.8	942	73

\* Values are mean (SE). RT, reaction time.

target word that rhymed but was orthographically dissimilar to one of the preceding two words (e.g., jazz, has). The dissimilar targets required the participant to make a judgment that was not based solely on overlapping letters.

**Visual control task**

The three stimuli were abstract, non-linguistic symbols consisting of straight lines. Participants determined whether the third stimulus (e.g., / /) was the same as either the first stimulus (e.g., / \) or the second stimulus (e.g., / /).

**Auditory meaning and rhyming tasks**

The auditory tasks employed a different list of words that the visual tasks, equated to them in word frequency and part of speech. During the auditory tasks, a white fixation cross (+) was presented during the presentation of the auditory stimuli. Participants were asked to fixate on the cross during the entire trial.

**Auditory control task**

The three stimuli were high, medium and low frequency non-linguistic pure tones. The tones were 600 msec in duration and contained a 100 msec linear fade in and a 100 msec linear fade out. Participants determined whether the third stimulus (e.g., 700 Hz) was the same as either the first (e.g., 500 Hz) or second stimuli (e.g., 300 Hz).

**Stimulus characteristics**

All tasks were structured in a similar way so patterns of brain activation could be directly compared across tasks. First, the tasks consisted of words with similar written word frequency for children and adults [Educator’s Word Frequency Guide, 1996] and similar adult word frequency for written and spoken language [Baayen et al., 1995]. Second, no homophones were included in the experimental lists. Third, the tasks contained about the same number of nouns (55–65%), verbs (25–35%) and adjectives (10–20%) based on their most frequent usage in the Oxford English Dictionary. Fourth, none of the words were more than two syllables in length.

**Experimental procedure**

**MRI practice session**

The participant was acclimated to the scanner environment in a simulator. The participant was slid on a mat into the open tube-like structure of the simulator. From this position, the participant was able to view a computer monitor about 40 cm directly above them. The participant then put on headphones and grasped a button box in the right hand. After the participant seemed comfortable, the participant practiced a full-length version of each experimental task (Table I). Different stimuli (matched in their stimulus characteristics) were used in the practice and fMRI sessions.

### **MRI data acquisition**

After screening, the participant was asked to lie down on the scanner bed. The head position was secured with a specially designed vacuum pillow (Bio-nix, Toledo, OH) that allowed for the insertion of two earphones (for the auditory sessions). An optical response box (Lightwave Medical, Burnaby, Canada) was placed in the participant's right hand and a compression alarm ball was placed in the left hand. The head coil was positioned over the participant's head and a goggle system for the visual presentation of stimuli (Avotec, Jensen Beach, FL) was secured to the head coil. Each imaging session took <1 hr.

All images were acquired using a 1.5 Tesla GE scanner. Gradient echo localizer images were acquired to determine the placement of the functional slices. For the functional imaging studies, a susceptibility weighted single-shot EPI (echo planar imaging) method with BOLD (blood oxygenation level-dependent) was used. The following scan parameters were used: TE = 40 msec, flip angle = 90°, matrix size = 64 × 64, field of view = 22 cm, slice thickness = 4 mm, number of slices = 32. These scanning parameters resulted in a 3.437 × 3.437 × 4 mm voxel size. The acquisition of this volume was repeated every 3 sec (TR = 3,000 msec) for a total of 9 min per run.

At the end of the functional imaging session, a high resolution, T1 weighted 3D image was acquired (SPGR, TR = 21 msec, TE = 8 msec, flip angle = 20°, matrix size = 256 × 256, field of view = 22 cm, slice thickness = 1 mm). These scanning parameters resulted in a 0.86 × 0.86 × 1 mm voxel size. The orientation of this 3D volume was identical to the functional slices.

### **Image data analysis**

Data analysis was carried out using SPM-99 (Statistical Parametric Mapping) for motion correction and statistical inference [Friston et al., 1994, 1995a,b].

The functional images are realigned to the last functional volume in the scanning session using affine transformations. No individual runs (rhyming and meaning in the visual or auditory modality) had more than 2.5 mm maximum displacement (less than the voxel size) from the beginning to the end of the run for any participant in the x-plane (median = 0.18, range = 0.02–0.70), y-plane (median = 0.34, range = 0.13–1.65), or z-plane (median = 0.57, range = 0.10–2.48). Furthermore, no individual runs had more than 3 degrees of maximum displacement in rotation from

the beginning to the end of the run for pitch (median = 0.89, range = 0.27–2.74), yaw (median = 0.61, range = 0.14–1.76), or roll (median = 0.51, range = 0.09–1.64). All statistical analyses were conducted on these movement-corrected images.

Images were then segmented and the gray-white matter information is used to co-register the structural and functional images. The co-registered images were normalized to the MNI stereotaxic template (12 linear affine parameters for brain size and position, 8 non-linear iterations and 2 × 2 × 2 nonlinear basis functions for subtle morphological differences). The MNI template used for normalization by SPM-99 is similar to the Talairach and Tournoux [1998] stereotaxic atlas [Talairach and Tournoux, 1988]. The major difference between these two atlases is in the inferior portion of the temporal lobes [Calder et al., 2000].

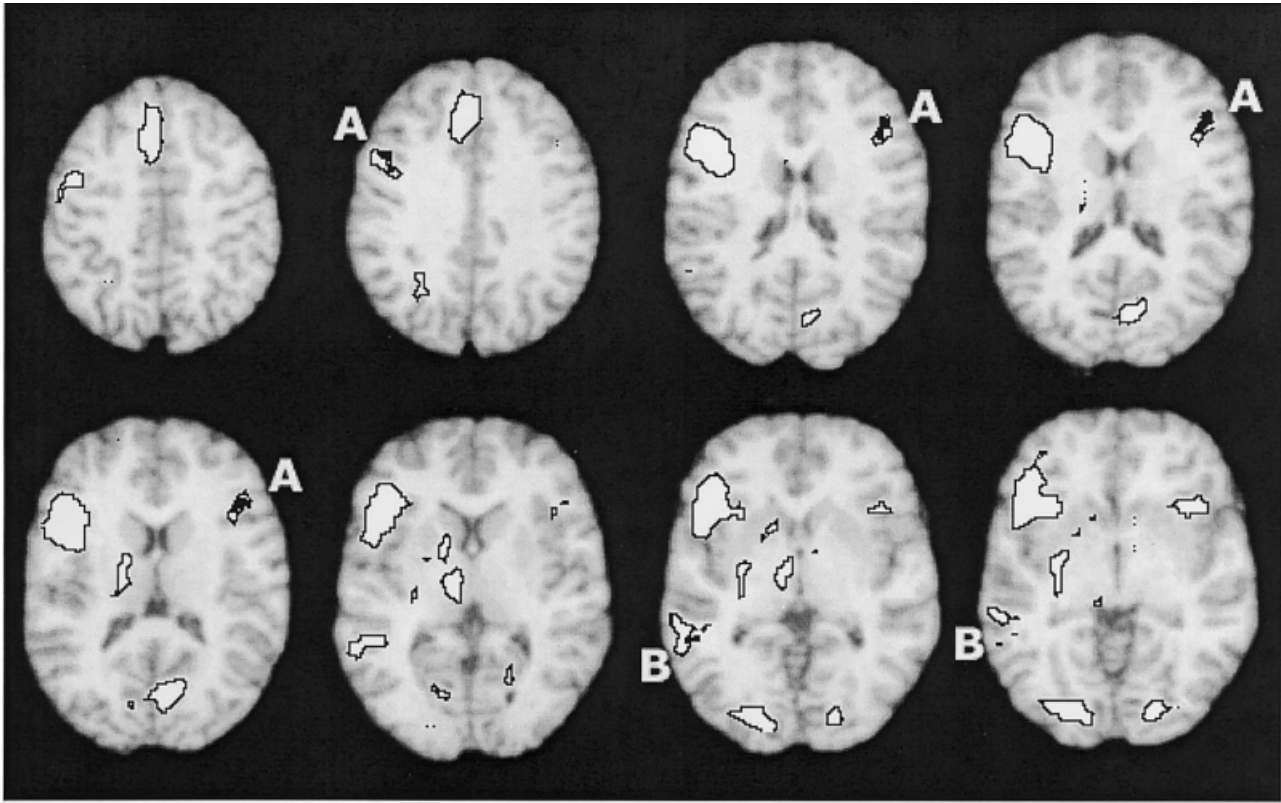
Statistical analyses were calculated on the smoothed data (7 mm isotropic Gaussian kernel) using a delayed boxcar design with a 6-sec delay from onset of block to account for the lag in hemo-dynamic response. A high pass filter was applied equal to 2 cycles of the experimental and control conditions (216 sec) to remove low frequency effects such as signal drift, cardiac and respiratory pulsations. We used global normalization to scale the mean of each scan to a common value to correct for whole brain differences over time.

Random effect statistics allow generalization to the population. In the first level analysis, we calculated parameter estimate images for individual subjects across the entire brain. For each individual, we calculated contrasts [experimental–control] to analyze the 2-word judgment tasks (meaning, rhyming) in the two modalities (visual, auditory). In the second level analysis, these parameter estimate images were entered into statistical analyses. A one-sample z-test compared each voxel across all participants to determine whether the activation during a condition was significant (i.e., significantly >0); a two-sample z-test was used to determine whether the magnitude of activation across conditions was significantly different. All reported areas of activation were significant using  $P < 0.001$  uncorrected for the voxel level and contained a cluster size greater than 11 voxels. We will concentrate on reporting the results for our regions of interest including unimodal visual regions (fusiform gyrus and surrounding area), unimodal auditory regions (superior temporal gyrus and surrounding area), middle temporal gyrus and inferior frontal gyrus. All areas of significant activation are presented in the tables.

**TABLE II. Significant activation for the meaning and rhyming tasks in the visual and auditory modality as compared to the control tasks\***

	Location			Significance		Coordinates			
	Area	H	BA	Z-test	Voxels	X	Y	Z	
Visual									
Meaning	Medial frontal gyrus	—	6	6.14	285	-3	36	36	
	Inferior frontal gyrus	L	46	6.26	631	-45	21	24	
	Cuneus	R	18	4.18	25	15	-72	15	
	Posterior cingulate	R	30	4.66	16	21	-57	6	
	Caudate	L	—	4.55	171	-15	15	3	
	Middle temporal gyrus	L	21	3.77	24	-60	-51	0	
	Middle occipital gyrus	R	18	4.64	45	21	-87	-3	
	Inferior occipital gyrus	L	18	5.36	137	-27	-90	-6	
	Putamen	L	—	3.76	37	-27	-3	-6	
	Inferior frontal gyrus	R	47	4.89	147	42	21	-9	
	Anterior cerebellar lobe	R	—	3.73	13	12	-54	-9	
	Posterior cerebellar lobe	R	—	3.65	16	36	-69	-24	
	Posterior cerebellar lobe	L	—	3.95	57	-36	-57	-27	
	Posterior cerebellar lobe	R	—	4.23	75	15	-78	-30	
	Rhyming	Posterior cerebellar lobe	—	—	3.82	16	0	-60	-33
Superior frontal gyrus		—	8	4.87	69	-3	21	51	
Medial frontal gyrus		—	8	4.33	57	0	39	39	
Angular gyrus		L	39	3.79	21	-27	-51	36	
Inferior frontal gyrus		L	45	5.35	657	-45	30	12	
Posterior cingulate		R	30	4.09	28	12	-63	12	
Middle occipital gyrus		L	19	4.42	99	-18	-90	-3	
Putamen		L	—	4.35	60	-30	-6	-3	
Inferior frontal gyrus		R	47	5.38	29	39	27	-9	
Middle occipital gyrus		R	19	3.82	21	30	-78	-9	
Fusiform gyrus		L	—	4.69	92	-45	-60	-21	
Posterior cerebellar lobe		R	—	4.70	49	9	-75	-30	
Auditory									
Meaning		Inferior frontal gyrus	L	45	4.76	203	-48	18	18
		Thalamus	L	—	4.04	28	-12	0	15
	Cuneus	L	18	4.75	101	-9	-78	6	
	Middle temporal gyrus	L	21	5.10	414	-51	-18	-6	
	Middle temporal gyrus	R	21	4.03	136	63	-15	-6	
	Inferior frontal gyrus	L	47	4.60	26	-33	21	-9	
	Anterior cerebellar lobe	L	—	4.41	23	-6	-33	-9	
	Inferior frontal gyrus	L	47	4.16	19	-45	36	-15	
	Posterior cerebellar lobe	R	—	5.42	103	12	-81	-27	
	Rhyming	Medial frontal gyrus	L	6	4.48	38	-6	36	36
Inferior frontal gyrus		L	45	4.35	110	-51	21	27	
Transverse temporal		R	41	4.48	13	42	-27	12	
Superior temporal gyrus		R	22	4.67	106	60	-9	-3	
Superior temporal gyrus		L	22	4.61	131	-54	-18	-3	
Fusiform gyrus		L	—	3.60	13	-45	-60	-18	
Posterior cerebellar lobe		R	—	4.50	42	15	-78	-24	

\* H, left (L) or right (R) hemisphere; BA, Brodmann's area of peak activation. Coordinates: -X left hemisphere, +X right hemisphere, -Y behind anterior commissure, +Y in front of anterior commissure, -Z below anterior-posterior commissure plane, +Z above anterior-posterior commissure plane. The coordinates represent the peak of activation in a cluster (number of voxels indicated) as determined by a z-test. Regions are sorted by Z coordinate.



**Figure 1.**

Activation maps of task differences for the visual modality. Letters label regions of interest. Solid black indicates areas of significantly more activation in the meaning than in the rhyming task (**A**:

inferior frontal gyrus; **B**: middle temporal gyrus). Black borders indicate areas of overlapping activation between the meaning and the rhyming tasks. The left side of the brain is on the left.

## RESULTS

### Behavioral performance

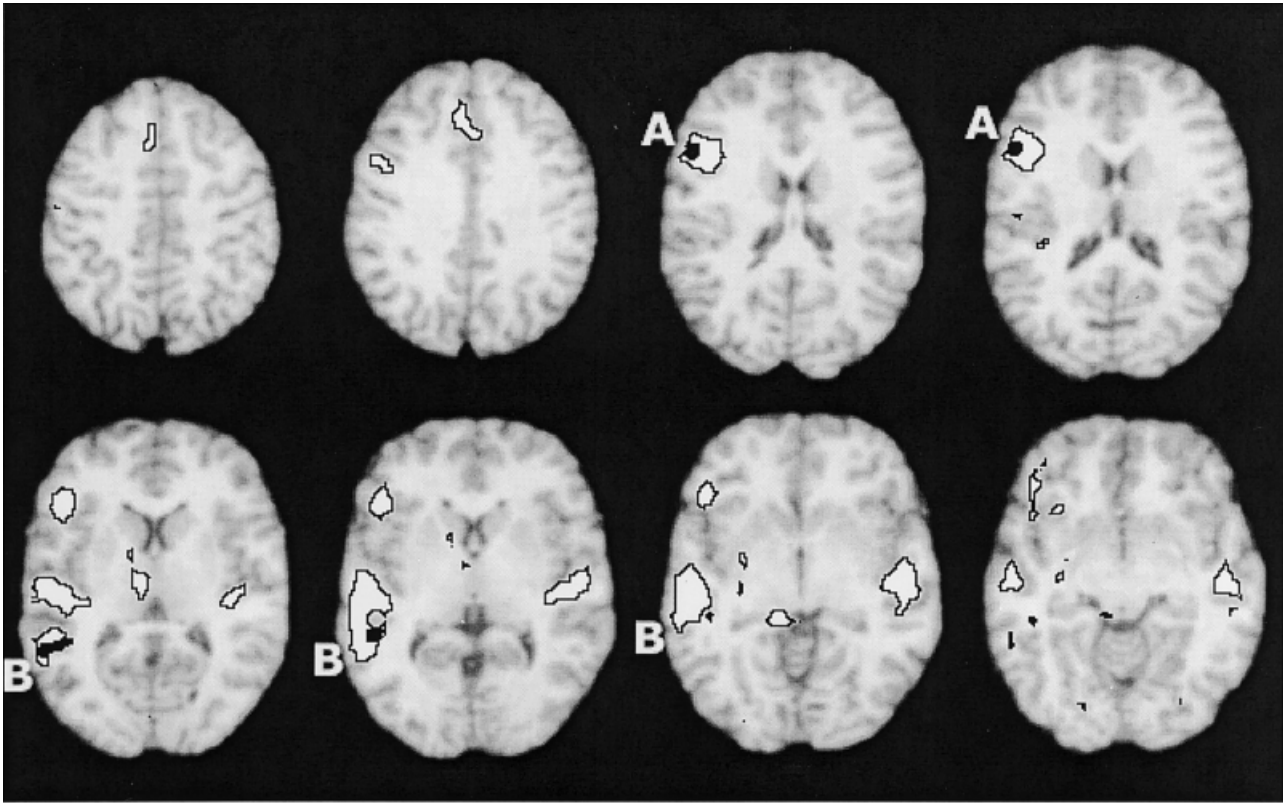
Table I presents behavioral data on the word judgment and control tasks. For the word judgment tasks, we calculated a task (meaning, rhyming) by modality (visual, auditory) by session (practice, test) ANOVA separately on accuracy and reaction time. This analysis showed that the meaning task had significantly slower reaction time than the rhyming task [ $F(1,95) = 4.69, P < 0.05$ ], and that the visual modality had significantly lower accuracy than the auditory modality [ $F(1,95) = 4.00, P < 0.05$ ]. We also calculated the same ANOVA as above for the control tasks. This analysis showed that the visual modality had significantly higher accuracy and faster reaction time than the auditory modality [ $F(1,95) = 5.66, P < 0.05$  for accuracy;  $F(1,95) = 22.49, P < 0.001$  for reaction time]. In addition, we calculated  $t$ -tests to determine if accuracy or reaction time differed between the word judgment

and control tasks. There was no significant difference in performance for accuracy, but reaction time for the control blocks was faster than for the experimental blocks,  $t(191) = 3.34, P < 0.01$ .

### Word judgment vs. control

A random effects model was used to examine differences in brain activation between the word judgment and control tasks. We calculated a total of four contrasts [experimental–control], one for each word judgment task (meaning and rhyming) in each of the two modalities (visual and auditory). Table II presents the significant results from the statistical comparison of each word judgment task vs. its control task.

Among unimodal visual regions, there was activation for both visual tasks in bilateral middle to inferior occipital gyrus (BA 18, 19) that extended in the left hemisphere to fusiform gyrus (BA 37). Among unimodal auditory regions, there was activation for both auditory tasks in bilateral superior temporal gyrus



**Figure 2.**

Activation maps of task differences for the auditory modality. Letters label regions of interest. Solid black indicates areas of significantly more activation in the meaning than in the rhyming task (**A**: inferior frontal gyrus; **B**: middle temporal gyrus). Black

borders indicate areas of overlapping activation between the meaning and the rhyming tasks. The left side of the brain is on the left.

(BA 22). For the auditory meaning task, this bilateral superior temporal activation extended into left and right middle temporal gyri (BA 21), although there was more activation in left (414 voxels) than in right temporal regions (136 voxels). There was also activation in left middle temporal gyrus for the visual meaning task. In terms of inferior frontal activation, both visual tasks tended to produce more activation in inferior frontal gyrus (BA 9, 46, 45) than the auditory tasks and all but the auditory rhyming task produced activation in the ventral portion of inferior frontal gyrus (BA 47).

#### Task-specific analyses within modality

A random effects model was used to examine differences between the meaning and rhyming tasks. We calculated the following analyses separately for each modality. We calculated meaning-control minus rhyming-control to show regions activated signifi-

cantly more in the meaning task than in the rhyming task. We calculated rhyming-control minus meaning-control activation maps to show regions activated significantly more in the rhyming task than in the meaning task. Finally, we examined overlap in activation between the meaning and rhyming tasks. This analysis produced a map of areas that were significantly activated in both the meaning and rhyming task, but that were not significantly different between the tasks.

Table II presents the data for the statistical comparison of the meaning task vs. the rhyming task separately for the visual (Fig. 1) and auditory modality (Fig. 2). The meaning tasks produced several peaks of activation when compared to the rhyming tasks. In particular, for the visual meaning task, there was activation in left middle temporal gyrus (BA 21), left inferior frontal gyrus (BA 9, 46, 45) and right inferior frontal gyrus (BA 46, 45). The results were similar for the auditory meaning task. In particular, there were peaks of activation in left inferior frontal gyrus (BA 46)

**TABLE III. Task specific activation for meaning and rhyming tasks for the visual and auditory modality\***

	Location			Significance		Coordinates		
	Area	H	BA	Z-test	Voxels	X	Y	Z
Visual								
ME–RH	Inferior frontal gyrus	L	9	4.22	22	–42	9	36
	Inferior frontal gyrus	R	46	4.29	27	51	33	18
	Middle temporal gyrus	L	21	4.08	36	–51	–45	–3
RH–ME	Posterior cingulate gyrus	R	23	3.92	15	12	–45	24
Auditory								
ME–RH	Inferior frontal gyrus	L	45	5.19	23	–51	21	18
	Cuneus	L	18	3.82	15	–21	–93	3
	Middle temporal gyrus	L	21	4.54	146	–45	–33	–6
	Posterior cerebellar lobe	R	—	3.68	33	24	–72	–12
	Posterior cerebellar lobe	R	—	3.39	13	12	–81	–30
RH–ME	Middle cingulate gyrus	—	31	3.38	15	–3	–27	42

\* ME, meaning; RH, rhyming; H, left (L) or right (R) hemisphere; BA, Brodmann's area of peak activation. Coordinates: –X left hemisphere, +X right hemisphere, –Y behind anterior commissure, +Y in front of anterior commissure, –Z below anterior-posterior commissure plane, +Z above anterior-posterior commissure plane. The coordinates represent the peak of activation in a cluster (number of voxels indicated) as determined by a Z-test. Regions are sorted by Z coordinate.

and in left middle temporal gyrus (BA 21) that extended into left superior temporal gyrus (BA 22). The rhyming tasks did not produce much task specific activation. The visual rhyming task produced activation in posterior cingulate gyrus (BA 23) and the auditory rhyming task produced activation in middle cingulate gyrus (BA 31).

### Modality-independent analyses

A random effects model was used to examine differences between the meaning and rhyming tasks independent of modality. In the first level of analysis, we calculated contrasts between the tasks within each modality (see above). In the cross-modal analysis, the parameter estimate images for each contrast from both modalities were entered into a one-sample z-test for the spelling or the rhyming task. This analysis produced a statistical map of areas that were activated by the task regardless of whether the words were heard or read.

Table IV and Figure 3 present the data for the statistical comparison of the meaning versus the rhyming task independent of modality. For the meaning task, there were large areas of activation in left inferior frontal gyrus (BA 46, 47), right inferior frontal gyrus (BA 46) and left middle temporal gyrus (BA 21). For the rhyming task, there was activation in the left supramarginal gyrus (BA 40), posterior cingulate gyrus (BA 23) and middle cingulate gyrus (BA 31).

For the modality independent analysis, each participant had two parameter estimate images (auditory and visual) entered into the two-sample z-test. This approach effectively doubled the sample size, thus increasing the sensitivity for detecting small-amplitude activation that might miss the statistical threshold used in the single modality analysis. Our modality independent analysis demonstrated activation in right inferior frontal gyrus that just missed the statistical criterion for activation in the visual within-modality analysis.

### DISCUSSION

By comparing meaning to rhyming judgments for both written and spoken words, our study was able to determine a set of brain areas subserving semantic processing independent of input modality. Our word comprehension tasks produced task-specific (Table III, Figs. 1,2) and modality-independent (Table IV, Fig. 3) activation in the inferior frontal and the middle temporal gyri. With respect to the frontal activations, the within-modality analyses showed that written word comprehension produced activation in the left (BA 9, 46, 45) and the right inferior frontal gyrus (BA 46, 45), whereas spoken word comprehension produced activation only in the left inferior frontal gyrus (BA 46). These regions of the frontal lobe at least partially overlap with the boundaries of the region that is generally included within Broca's area. In general, there



**TABLE IV. Task specific activation for meaning and rhyming tasks independent of modality\***

	Location			Significance		Coordinates		
	Area	H	BA	Z-test	Voxels	X	Y	Z
ME–RH	Anterior cingulate gyrus	L	32	3.65	22	–6	18	42
	Inferior frontal gyrus	L	46	5.35	80	–54	21	18
	Inferior frontal gyrus	R	46	4.74	60	48	30	18
	Cuneus	R	23	4.03	70	12	–81	12
	Lingual gyrus	L	19	3.78	19	–27	–63	0
	Lingual gyrus	R	18	3.48	12	18	–84	–3
	Middle temporal gyrus	L	21	4.69	275	–48	–33	–6
	Inferior frontal gyrus	L	47	4.03	21	–48	21	–9
RH–ME	Middle cingulate gyrus	L	31	3.66	43	–6	–36	45
	Supramarginal gyrus	L	40	3.59	39	–57	–21	39
	Posterior cingulate gyrus	—	23	4.11	33	0	–39	24
	Precuneus	L	31	3.56	27	–6	–63	24
	Superior frontal gyrus	L	10	3.46	30	–9	57	–12

\* H, left (L) or right (R) hemisphere; BA, Brodmann’s area of peak activation. Coordinates: –X left hemisphere, +X right hemisphere, –Y behind anterior commissure, +Y in front of anterior commissure, –Z below anterior-posterior commissure plane, +Z above anterior-posterior commissure plane. The coordinates represent the peak of activation in a cluster (number of voxels indicated) as determined by a z-test. Regions are sorted by Z coordinate.

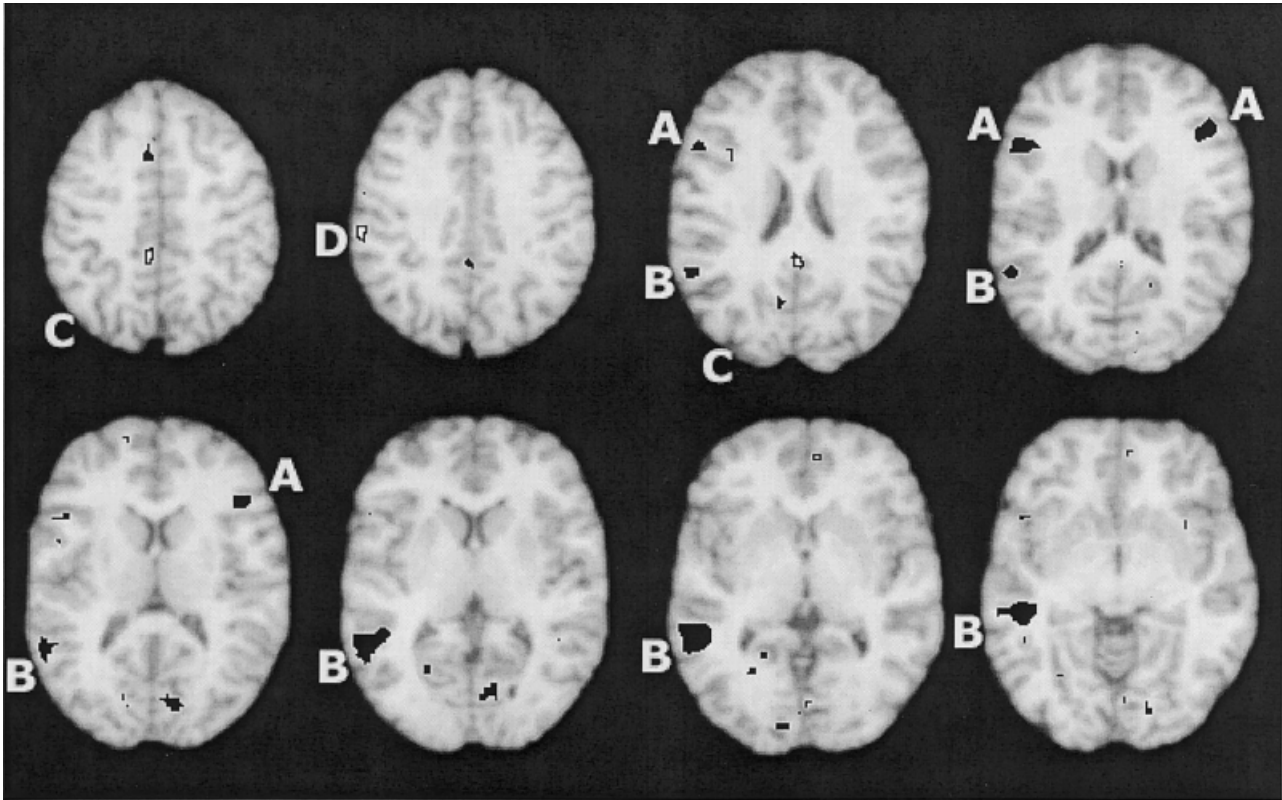
was greater modality-independent activation in the left than in the right inferior frontal gyrus. This asymmetry is consistent with the well-known left hemisphere dominance of language processing and with other imaging studies based on semantic tasks [Binder et al., 1997; Illes et al., 1999; Kapur et al., 1996].

The region of the middle temporal gyrus where we identified activations contains heteromodal association cortex and can be included with the complex of areas constituting Wernicke area [Mesulam, 1998]. The task-specific (Table III, Figs. 1,2) and modality-independent (Table IV, Fig. 3) activations we observed in the left middle temporal gyrus (BA 21) are consistent with other studies that have examined semantic processing along different dimensions including judging whether a word is abstract or concrete, living or non-living and determining its category [Friederici et al., 2000; Price et al., 1997; Pugh et al., 1996]. In keeping with the results of Chee et al. [1999], our study showed greater activation in the left middle temporal gyrus for the auditory than for the visual meaning task suggesting that this region may be more sensitive to semantic processing of spoken rather than written words.

Discrepancies in the existing literature on the functional anatomy of semantic processing could be attributed to differences in the nature of the semantic and baseline tasks used in the individual experiments [Chee et al., 1999; Mummery et al., 1998; Perani et al., 1999; Petersen et al., 1988, 1989; Seger

et al., 2000]. In terms of the baseline task, our study required participants to determine whether two words rhymed (e.g., bank, tank). In contrast, other studies examining modality differences required participants to judge the case of written words. Such studies cannot separate the activation associated with word comprehension from the activation associated with word form recognition. Previous studies have also used baseline conditions such as reading aloud or auditory word repetition that might have engaged some semantic processing, leading to an underestimation of the brain regions involved in word comprehension. Another important feature of our study was the equivalent difficulty level of the semantic and rhyming tasks.

We had proposed a theoretical model of language comprehension according to which visual and auditory word forms are encoded in modality-specific association areas (fusiform gyrus for written words and superior temporal gyrus for spoken words) before being relayed to heteromodal cortices in Broca and Wernicke areas. These heteromodal areas provide critical processing nodes for linking modality-specific word forms with the distributed associations that give them meaning [Mesulam, 1998]. Our findings support this model. They show that spoken input activated auditory association cortex in the superior temporal gyrus, that written input activated the visual association cortex in the fusiform gyrus and that the semantic



**Figure 3.**

Activation maps for task differences that are independent of modality. Letters label regions of interest. Solid black indicates areas of significantly more activation in the meaning than in the rhyming task. (A: inferior frontal gyrus; B: middle temporal gyrus). Black

borders indicate areas of significantly more activation in the rhyming than in the meaning task. (C: middle to posterior cingulate gyrus; D: supramarginal gyrus). The left side of the brain is on the left.

integration of word forms in either modality activated heteromodal regions in the inferior frontal gyrus and the middle temporal gyrus. Our results also show that these heteromodal areas of the language network displayed a leftward asymmetry of activation that is consistent with the well-known dominance of the left hemisphere for language.

#### ACKNOWLEDGMENT

We thank S. Brennan, Y. Harasaki and F. Van Santen for their assistance in stimulus development, and K. Bettenhausen, J. Rex and C. Wolf for their assistance in conducting the behavioral study. We thank N. Christian, P. Springer, and R. Salzman for their operation of the MRI. We also thank the students, teachers, and administrators at Pope John XXIII School, Saint Athanasius School, and Saint Peter's Catholic School for their participation.

#### REFERENCES

- Baayen RH, Piepenbrock R, Gulikers L (1995): The CELEX Lexical Database (Version Release 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T (1997): Human brain language areas identified in functional magnetic resonance imaging. *J Neurosci* 17:353–362.
- Calder AJ, Lawrence AD, Young AW (2001): Neuropsychology of fear and loathing. *Nat Rev Neurosci* 2:352–363.
- Chee MWL, O'Craven KM, Bergida R, Rosen BR, Savoy RL (1999): Auditory and visual word processing studied with fMRI. *Hum Brain Mapp* 7:15–28.
- Demonet J-F, Chollet F, Ramsay T, Cardebat D, Nespoulous J-L, Wise R, Rascoul A, Frackowiak R (1992): The anatomy of phonological and semantic processing in normal subjects. *Brain* 115:1753–1768.
- Duncan J, Seitz RJ, Kolodny J, Bor D, Herzog H, Ahmed A, Newell FN, Emslie H (2000): A neural basis for general intelligence. *Science* 289:457–460.
- Educator's Word Frequency Guide (1996): Brewster, NY: Touchstone Applied Science Associates, Inc.

- Friederici AD, Opitz B, Cramon DYv (2000): Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types. *Cereb Cortex* 10:698–705.
- Friston KJ, Ashburner J, Frith CD, Poline JB, Heather JD, Frackowiak RSJ (1995a): Spatial registration and normalization of images. *Hum Brain Mapp* 2:1–25.
- Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SCR, Frackowiak RSJ, Turner R (1995b): Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2:189–210.
- Friston KJ, Jezzard P, Turner R (1994): Analysis of functional MRI times-series. *Hum Brain Mapp* 1:153–171.
- Illes J, Francis WS, Desmond JE, Gabrieli JDE, Glover GH, Poldrack R, Lee CJ, Wagner AD (1999): Convergent cortical representation of semantic processing in bilinguals. *Brain Lang* 70:347–363.
- Kapur S, Tulving E, Cabeza R, McIntosh AR, Houle S, Craik FI (1996): The neural correlates of intentional learning of verbal materials: a PET study in humans. *Cogn Brain Res* 4:243–249.
- Mesulam MM (1998): From sensation to cognition. *Brain* 121:1013–1052.
- Mummery CJ, Patterson K, Hodges JR, Price CJ (1998): Functional neuroanatomy of the semantic system: divisible by what? *J Cogn Neurosci* 10:766–777.
- Nelson DL, McEvoy CL, Schreiber TA (1998): The University of South Florida word association, rhyme, and word fragment norms. <http://www.usf.edu/FreeAssociation/>
- Perani D, Schnur T, Tettamanti M, Gorno-Tempini M, Cappa SF, Fazio F (1999): Word and picture matching: a PET study of semantic category effects. *Neuropsychologia* 37:293–306.
- Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME (1988): Positron emission tomographic studies of the cortical anatomy of single word processing. *Nature* 331:585–589.
- Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME (1989): Positron emission tomographic studies of the processing of single words. *J Cogn Neurosci* 1:153–170.
- Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10:15–35.
- Price CJ, Moore CJ, Humphreys GW, Wise RJS (1997): Segregating semantic from phonological processes during reading. *J Cogn Neurosci* 9:727–733.
- Pugh KR, Shaywitz BA, Shaywitz SE, Constable RT, Skudlarski P, Fulbright RK, Bronen RA, Shankweiler DP, Katz L, Fletcher JM, Gore JC (1996): Cerebral organization of component processes in reading. *Brain* 119:1221–1238.
- Seger CA, Desmond JE, Glover GH, Gabrieli JD (2000): Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships. *Neuropsychology* 14:361–369.
- Talairach J, Tournoux P (1988): Co-planar stereotaxic atlas of the human brain. New York: Thieme.
- Wagner AD, Schacter DL, Rotte M, Koutstaal W, Maril A, Dale AM, Rosen BR, Buckner RL (1998): Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* 281:1188–1191.