

Imagery in sentence comprehension: an fMRI study

Marcel Adam Just,* Sharlene D. Newman, Timothy A. Keller,
Alice McEleney, and Patricia A. Carpenter

Center for Cognitive Brain Imaging, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received 2 April 2003; revised 28 August 2003; accepted 29 August 2003

This study examined brain activation while participants read or listened to high-imagery sentences like *The number eight when rotated 90 degrees looks like a pair of spectacles* or low-imagery sentences, and judged them as true or false. The sentence imagery manipulation affected the activation in regions (particularly, the intraparietal sulcus) that activate in other mental imagery tasks, such as mental rotation. Both the auditory and visual presentation experiments indicated activation of the intraparietal sulcus area in the high-imagery condition, suggesting a common neural substrate for language-evoked imagery that is independent of the input modality. In addition to exhibiting greater activation levels during the processing of high-imagery sentences, the left intraparietal sulcus also showed greater functional connectivity in this condition with other cortical regions, particularly language processing regions, regardless of the input modality. The comprehension of abstract, nonimaginal information in low-imagery sentences was accompanied by additional activation in regions in the left superior and middle temporal areas associated with the retrieval and processing of semantic and world knowledge. In addition to exhibiting greater activation levels during the processing of low-imagery sentences, this left temporal region also revealed greater functional connectivity in this condition with other left hemisphere language processing regions and with prefrontal regions, regardless of the input modality. The findings indicate that sentence comprehension can activate additional cortical regions that process information that is not specifically linguistic but varies with the information content of the sentence (such as visual or abstract information). In particular, the left intraparietal sulcus area appears to be centrally involved in processing the visual imagery that a sentence can evoke, while activating in synchrony with some core language processing regions.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Mental imagery; Sentence processing; fMRI

Introduction

Many types of thinking, particularly language comprehension, entail the use of mental imagery. Understanding a text on architecture or automobile design seems impossible without mental imagery.

* Corresponding author. Center for Cognitive Brain Imaging, Department of Psychology, Carnegie Mellon University, Baker Hall, Pittsburgh, PA 15213. Fax: +1-412-268-2804.

E-mail address: just@cmu.edu (M.A. Just).

Available online on ScienceDirect (www.sciencedirect.com).

Such multimodal thinking requires that the activity of several cortical systems be coordinated, such that the meaning of a sentence can be extracted by a language understanding system and the sentence content can be used to construct a mental image of the objects to which the sentence refers. For example, to evaluate a sentence like *The number eight when rotated 90 degrees looks like a pair of spectacles*, a reader must process the content of the sentence, retrieve a mental image of an eight-shape, mentally apply a rotation transformation to it, and then evaluate the resulting image.

Previous studies have indicated that mental imagery generated from verbal instructions and from visual encoding activate similar cortical regions (Mazoyer et al., 2002; Mellet et al., 1996, 1998, 2002). For example, a study conducted by Mellet et al. (1998) compared the cortical activation when participants listened to definitions of concrete, highly imaginable nouns vs. abstract nouns. The concrete nouns elicited greater activation in the prefrontal regions related to working memory and in the ventral areas related to object processing, whereas the abstract task elicited greater activation in temporal regions. Although several studies examining mental imagery have observed activation of the parietal area (Dieber et al., 1998; Ishai et al., 2000; Just et al., 2001; Kosslyn et al., 1993, 1996, 1999; Mellet et al., 1996, 2000), the concrete noun task of Mellet et al. (1998) failed to observe such activation. This lack of parietal involvement specific to the imagery related to the concrete noun task is interesting because this region has been shown to collaborate closely with prefrontal regions in the performance of working memory processes. In fact, it has been suggested that “parietal and frontal regions may mediate the retrieval of object representations from long-term memory, their maintenance in a working memory buffer, and the attention required to generate those mental images” (Ishai et al., 2000).

There is also a possibility that the neural activity that underlies the imagery in language processing is affected by the presentation modality of the language (i.e., written vs. spoken). For example, the neural activity elicited in primary visual cortex during mental imagery following verbal vs. visual encoding was different (Mellet et al., 2000); there was less primary visual activation during imagery after visual encoding compared to after verbal encoding, suggesting that presentation modality may indeed affect later imagery processing. Another study by Eddy and Glass (1981) examined how the visual processes in reading might be related to the visual imagery processes that a sentence engenders, comparing visual and auditory sentence presentation modes. The high-imag-

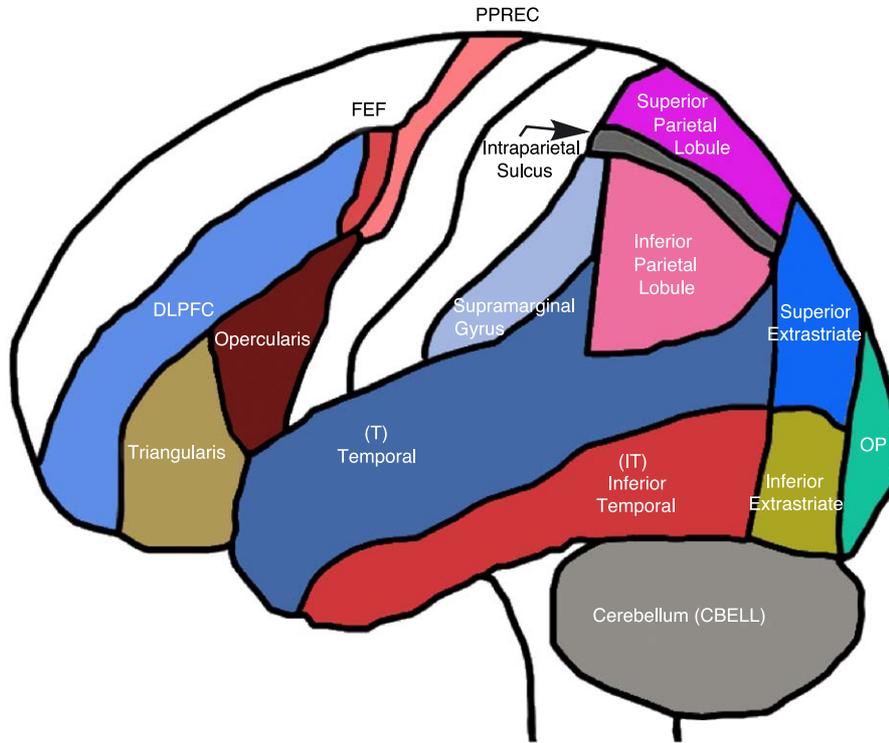


Fig. 1. Schematic depiction of regions of interest.

ery sentences took longer to verify as true or false than low-imagery sentences when the sentences were presented visually, but not when they were presented auditorily. These findings again suggest that the presentation modality of a sentence may affect the processing of the subsequent imagery.

The current study examined mental imagery processes in the context of a language comprehension task. One of the main goals was to examine the interaction between two somewhat separable neural systems, the mental imagery and language processing systems. To accomplish this goal, we used fMRI to measure not only the activation levels but also the functional connectivities of the regions believed to be involved in mental imagery, to determine the relations between the two systems. A second goal was to

examine the effect of the input modality, comparing the effect on the imagery-related activation when the sentences were either heard or read. To accomplish this goal, we compared the comprehension of high- and low-imagery sentences in both the visual and auditory modalities.

General methods

MRI acquisition parameters

The images were acquired on a GE 3.0 Tesla scanner with commercial birdcage, quadrature-drive radio frequency coils at the

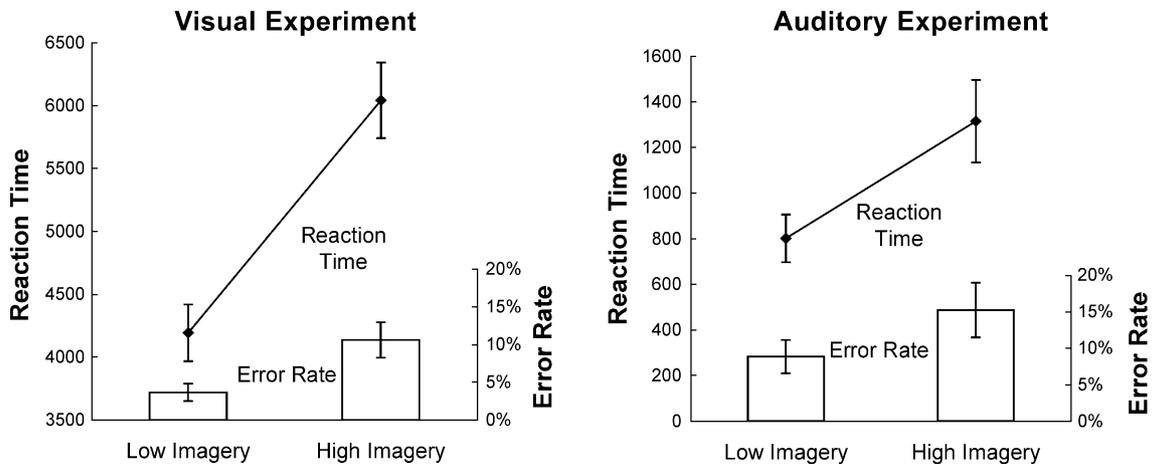


Fig. 2. Behavioral results from both the visual and auditory experiments. Error rates are shown in the bar graphs and response times in line graphs. The error bars indicate the standard error of the mean across participants.

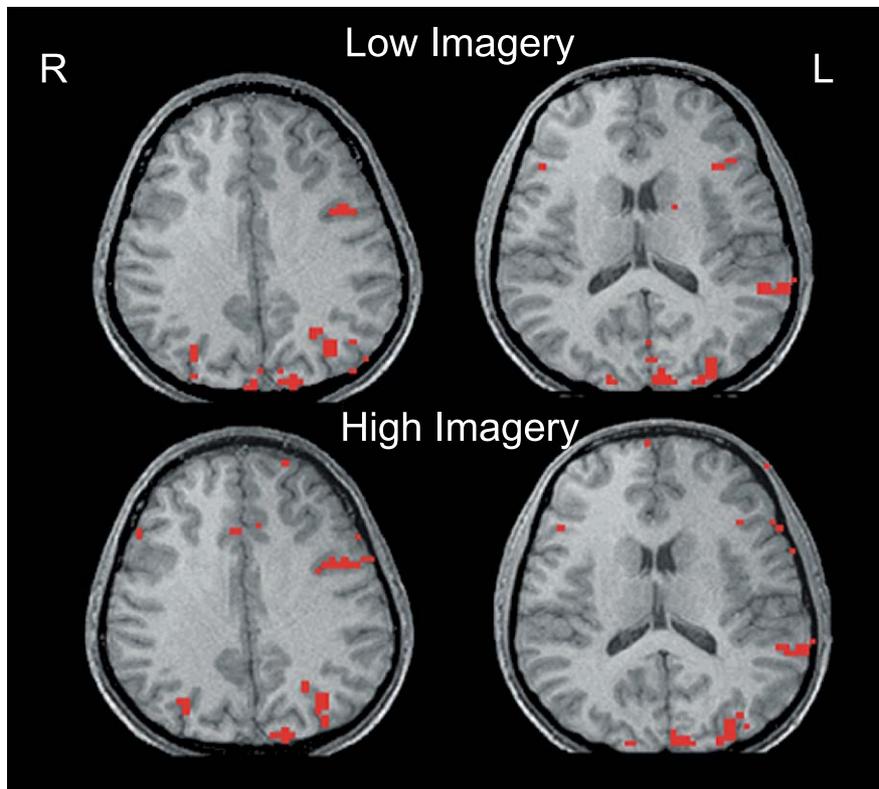


Fig. 3. Visual experiment activation map. Activation from two axial slices (the right figure is more superior) of a typical participant (activation level approximately equal to mean). The top two slices show activation resulting from the low-imagery condition and the bottom two the high-imagery condition.

Magnetic Resonance Research Center of the University of Pittsburgh Medical Center. The functional images were acquired using a gradient echo, resonant echo planar pulse sequence, with fourteen 5-mm-thick axial slices using the following parameters: TR = 3000 ms, TE = 25 ms, flip angle = 90° , voxel size = $3.125 \times 3.125 \times 5$ mm with a 1-mm gap. The means of the images corresponding to each of the 14 functional slices were registered to a high-resolution, T1-weighted structural volume scan of each participant. The volume scan was constructed from 124 three-dimensional SPGR axial images collected with TR = 25 ms, TE = 4 ms, flip angle = 40° , FOV = 24×18 cm, resulting in voxel dimensions of $0.9375 \times 0.9375 \times 1.5$ mm.

fMRI data analysis

Image preprocessing (including de-ghosting, mean correction, motion correction, and trend correction) was performed using FIASCO (Eddy et al., 1996; Lazar et al., 2001; further description and tools are available at www.stat.cmu.edu/~fiasco/). To examine the condition effects on the level of the activation in each cortical region, it was important to use an a priori, independent method of defining the ROIs. Anatomically defined ROIs were drawn on the structural images, without the activation overlaid, for each participant using the parcellation scheme described by Rademacher et al. (1992) and further refined by Caviness et al. (1996). The surfaces of the regions of interest are shown in Fig. 1. The method has been consistently and successfully used in many previous studies from several laboratories to compare the amount of activation in a given region across conditions (for further details, see Just et al., 2001).

fMRI-measured activation was quantified using FIASCO by first constructing a t map by computing the difference between each voxel's activation in each experimental condition and the baseline condition. Voxels whose signal change exceeded baseline by an amount determined by t tests with $t > 5.5$ were considered active. This high t threshold provides approximately similar or more conservative correction for multiple comparisons as compared to a Bonferroni correction with an alpha level of 0.01 given approximately 5000 voxels per participant in all regions of interest.

Functional connectivity analysis

If the activation in a set of cortical areas is highly synchronized, those regions are considered to be in close collaboration. One technique, *functional connectivity* or *effective connectivity* analysis (Buchel et al., 1999; Friston, 1994; Horwitz et al., 1998) assesses synchronization by measuring the correlation or covariance of the activation levels in two activated areas over some time period during which a cognitive activity is occurring, and examines how the synchronization between areas is modulated by a various determinants. The synchronization is taken as evidence of collaboration. In the current study, the functional connectivity between ROIs was assessed by computing the correlation of the activation time series between activated voxels in pairs of ROIs. Functionally connected areas might be collaborating or communicating, such that their activation levels are being modulated by partially overlapping workload. The correlations were based on only those periods when the task was being performed (thus excluding the fixation periods),

so that the time series indicates the momentary fluctuations in activation level during the task performance. The correlation between the averaged time courses from the activated voxels in each pair of ROIs was computed for each participant.

Visual experiment

Method

Participants

The participants were 12 right-handed college students (five females) aged between 18 and 30 years. Each participant gave signed informed consent approved by the University of Pittsburgh and Carnegie Mellon University Institutional Review Boards. Data from other participants, which included excessive head motion (usually indicated by having a maximum displacement of more than 0.3 voxels), were discarded.

Procedure

Participants were instructed to respond true or false to the presented sentences. Half of the sentences were high-imagery sentences that required mental imagery for their verification. They included sentences requiring mental rotation (e.g., *The letter W rotated 180 degrees and written in lower case, looks like the letter m*), evaluation of spatial relations (e.g., *On a map, Nevada is to the right of California*), combination of shapes (e.g., *The number nine can be constructed from a circle and a horizontal line, an example for which the correct response is “false”*), and comparison of visual aspects of common objects (e.g., *In terms of diameter, a quarter is larger than a nickel, which is larger than a dime*). The low-imagery sentences could be verified by referring to general knowledge, without the use of imagery (e.g., *Although now a sport, marathons started with Greek messengers bringing news*). The high- and low-imagery sentences were of approximately equal length. There were three epochs containing five high-imagery sentences each and three epochs of six low-imagery sentences each, for a total of 33 stimulus sentences, with the two types of epochs alternating. There was an extra sentence in each low-imagery epoch to keep the epoch lengths approximately equal in the two conditions. Pretesting showed that reading times were shorter for the low-imagery sentences. In this way an equal number of images were collected for each condition.

Participants were told “There will be some trials in which you will have to visualize the sentence in order to answer. For example: *The capital letter s with a vertical line through it looks like a dollar sign.*” (The stimulus sentences appear in the Appendix A.) In the debriefings after each scan, all of the participants who spontaneously commented on differential strategies (about one third of the participants) noted that they “visualized” or “built a mental picture” to perform the high-imagery condition.

Successive epochs were separated by 6-s rest periods during which participants were instructed to fixate on an asterisk displayed in the center of the screen. There were also four presentations of 24-s fixation epochs distributed across the session to provide a baseline measure of each participant’s brain activation, one at the beginning, one at the end, and two evenly dispersed between. Data from the 6-s rest intervals as well as the first 6 s of each experimental epoch were not included in the analysis to accommodate the hemodynamic onset and offset. The sentence and response options “true” and “false” were rear projected onto a

screen 20 cm from the participant. The true–false responses were made with two handheld response buttons. When a button was pressed, the displayed sentence was removed and the next sentence was presented. The participants were familiarized with the task before performing it in the scanner.

Results

Behavioral results

The accuracy rates were high for both the low- and the high-imagery sentences (96% and 91%, respectively) and not reliably different (see Fig. 2). Participants took longer on average to respond to the high-imagery sentences than the low-imagery sentences [5942 vs. 4121 ms, $F(1,11) = 106.66$, $P < 0.01$].

fMRI results overview

The high-imagery sentences produced more activation than the low-imagery sentences in the parietal lobule (left intraparietal sulcus) and also in the prefrontal region (pars opercularis, bilaterally). Fig. 3 shows the activation in two slices of one participant. The low-imagery sentences unexpectedly produced more activation than the high-imagery sentences in primary visual cortex and in left temporal cortex. (We use the term “temporal cortex” here and elsewhere to refer to the ROI consisting of the superior and middle temporal gyrus.)

Spatial processing. The cortical regions related to spatial processing were expected to produce more activation in the comprehension of high-imagery sentences compared to low-imagery sentences. The region in and around the left intraparietal sulcus and the left frontal eye fields (the anterior bank of the precentral sulcus area that bounds the middle temporal gyrus) both produced reliably more activation in the high-imagery condition than the low-imagery condition (see Table 1 and Fig. 4). This imagery effect was observed with respect to additional activation volume and percentage of change in signal intensity.

Language processing. The two subregions of Broca’s area (pars opercularis and pars triangularis) produced a differential response to the imagery manipulation. In both left and right pars opercularis, reliably more voxels were activated during comprehension of high-imagery sentences, as shown in Table 1 and Figs. 4 and 5. (The increase in the signal change measure was reliable only in the right hemisphere ROI.) In contrast to pars opercularis, pars triangularis revealed no effect of imagery in either hemisphere.

Both Broca’s area and left temporal cortex have been implicated in language processing. In the current study, however, there was a suggestion that the left temporal region responded differently than pars triangularis to the imagery manipulation. In the left temporal region, the percentage of change in signal intensity (but not the voxel count) was reliably *greater in the low-imagery condition*. The activation in this ROI was in the posterior portion of the temporal cortex, typically in the superior temporal gyrus but often extending into the superior temporal sulcus and sometimes into the adjoining portion of the middle temporal gyrus.

Primary sensory processing. Primary visual cortex, the calcarine sulcus ROI, revealed a response similar to that found in the left temporal region. The activation observed in primary visual cortex

Table 1
Activation levels and coordinates

Visual										Auditory											
Condition	Voxel count			Mean signal change (%)						x	y	z	Voxel count			Mean signal change (%)					
	Low-imagery	High-imagery	F(1,11)	Low-imagery	High-imagery	F(1,11)	Low-imagery	High-imagery	F(1,9)				Low-imagery	High-imagery	F(1,9)	Low-imagery	High-imagery	F(1,9)	x	y	z
<i>Parietal cortex</i>																					
L. intraparietal sulcus	5.58	9.92	10.05*	2.18	2.77	11.22*	-27	-63	43	0.5	9.5	5.85*	0.66	2.53	47.56*	-29	59	46			
L. superior parietal	1.08	2	2.27	2.23	2.98	1.6	-16	-66	53	0.3	2.8	8.21*	0.54	2.67	31.35*	-15	-64	53			
L. inferior parietal	1.17	1.92	1.76	2.23	2.59	1.19	-39	-56	36	2.5	4.4	6.14*	1.7	2.43	1.65	-46	-52	30			
R. intraparietal sulcus	3.5	5.25	4.2	2.12	2.53	1.6	29	-63	41	0.1	2.1	30*	0.58	2.16	83.47*	28	-53	46			
R. superior parietal	1.83	1.42	1.21	2.37	2.33	0.01	18	-65	48	0.1	1	3.49	0.54	2.09	13.29*	15	-63	54			
R. inferior parietal	0.17	0.92	2.05	1.04	2.16	3.71	42	52	37	6	4.8	1.45	2.18	2.09	0.22	47	35	19			
<i>Frontal cortex</i>																					
L. frontal eye fields	1.5	3.5	10.15*	2.17	2.69	13.18*	-39	2	44	5.4	10.1	45.08*	2.12	2.9	46.19*	-41	4	42			
L. pars opercularis	3.17	6.67	8.74*	2.23	2.67	3.27	-44	11	27	4.9	7.3	2.37	2.42	2.46	0.02	-43	13	25			
L. pars triangularis	1.5	1.42	0.08	2.71	2.43	1.87	-42	26	14	3.3	3.8	0.51	2.39	2.04	0.5	-39	26	15			
L. DLPFC	1.75	3.08	2.69	2.69	2.6	0.12	-40	24	25	1.6	4	9.19*	1.56	2.39	7.41*	-34	21	38			
R. frontal eye fields	0.83	1.83	2.13	2.88	2.77	0.1	43	3	43	0.5	2.6	5.76*	1.36	2.45	20.31*	42	5	44			
R. pars opercularis	0.58	1.92	5.03*	1.7	2.84	7.17*	43	10	30	1.1	1.9	0.85	1.98	1.94	0.01	39	17	22			
R. pars triangularis	0.17	0.67	1.32	1.39	2.48	1.78	40	24	13	1	0.5	1.1	1.93	1.81	0.11	36	25	16			
R. DLPFC	0.83	1.42	2.37	1.36	2.27	1.59	38	32	27	0.8	1.9	2.23	1.6	2.25	7.19*	32	11	48			
<i>Temporal cortex</i>																					
L. temporal	7.42	7	0.11	2.99	2.48	9.45*	-53	-31	5	26.3	20.7	1.94	2.62	2.25	7.57*	-51	-29	8			
L. Heschl's	0	0.08	1	1.03	2.49	NA	-44	-18	12	10.8	9.4	1.11	2.82	2.44	10.15*	-49	-17	12			
R. temporal	0.5	0.42	0.06	1.75	1.67	0.04	48	-23	1	21.4	14.8	4.14	2.58	2.24	12.09*	53	-20	7			
R. Heschl's	0	0	NA	0	0	NA	NA	NA	NA	10.5	8.4	2.19	2.75	2.36	9.8*	49	-17	15			
<i>Occipital cortex</i>																					
Calcarine	32	18.27	8.37*	3.52	2.5	32.51*	0	-70	10	2.7	0.2	4.59	2.58	1.5	50.05*	1	-67	9			
L. extrastriate	12.58	15.33	1.63	2.9	2.92	0.02	-30	-64	-5	1	2.7	7.63*	1.18	2.55	11.96*	-44	-50	-2			
R. extrastriate	7.75	8.5	0.32	2.59	2.74	0.21	29	-64	-4	0.6	1.2	2.25	1.86	2.95	2.27	45	-52	1			

* $P < 0.05$; bold, italicized coordinates indicate significant modality differences at $P < 0.01$.

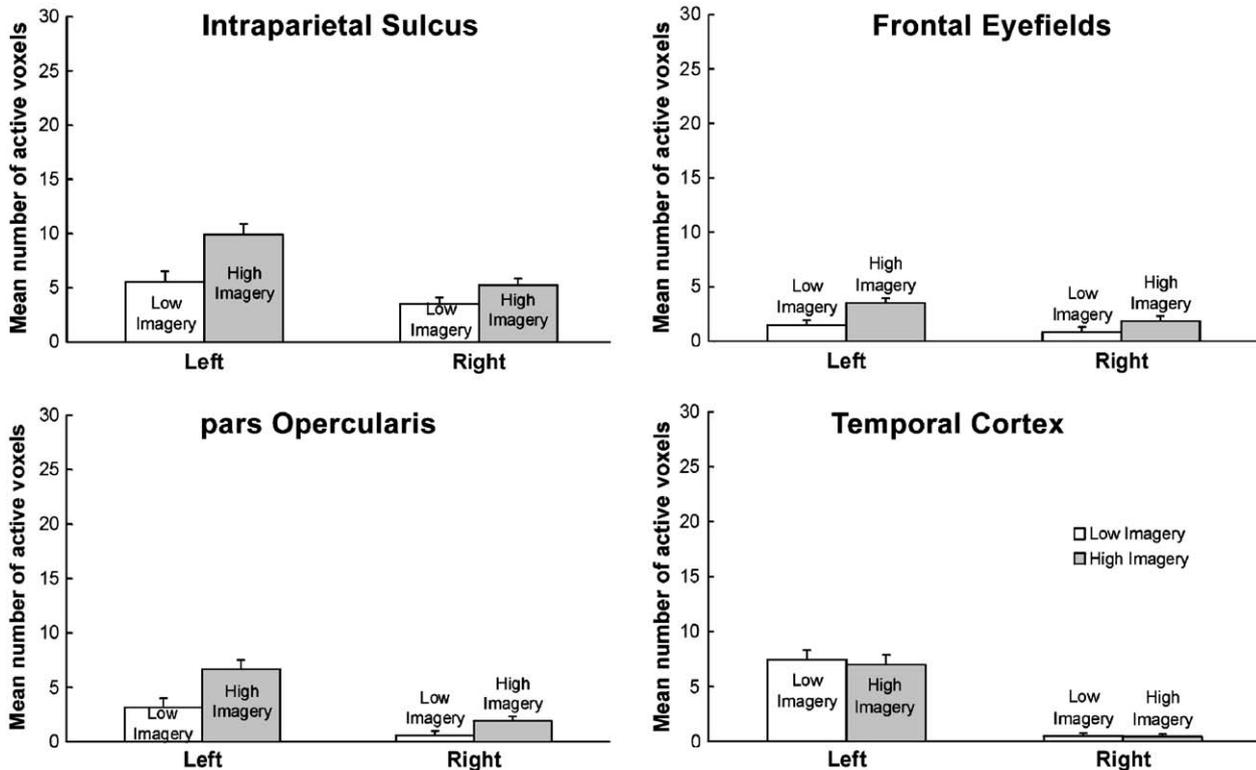


Fig. 4. Visual experiment voxel counts. The mean number of activated voxels from the intraparietal sulcus, frontal eye fields, pars opercularis, and temporal cortex. Error bars represent 95% confidence intervals based on the pooled *Mse* from the corresponding ANOVA (Loftus and Mason, 1994).

was reliably greater for the low-imagery condition compared to the high-imagery condition.

Laterality effects. As expected, this language task generally activated left hemisphere more so than the right. Three regions revealed reliably more activated voxels in the left hemisphere compared to the right. They included the intraparietal sulcus [$F(1,11) = 7.42, P < 0.05$], frontal operculum [$F(1,11) = 24.26, P < 0.01$], and the temporal area [$F(1,11) = 12.22, P < 0.01$].

Time courses. Although this was a block design study, it is useful to note that the time courses of the activation (in terms of percentage of change in signal intensity) were systematic and reflected the main results. To illustrate the time courses, Fig. 5 presents them for two of the key ROIs, LIPS and LT, indicating the differential effect of sentence imagery on these two regions.

Functional connectivity analysis. Functional connectivity analysis revealed differential effects of imagery on the left intraparietal sulcus area and the left temporal cortex. These two areas are proposed to be respectively involved in the imagery processing of high-imagery sentences and the semantic processing of low-imagery sentences. We measured the functional connectivities of these two areas to five left hemisphere ROIs, including ROIs centrally involved in language processing (pars opercularis, pars triangularis, dorsolateral prefrontal cortex, frontal eye fields, and inferior parietal), and compared them using an ANOVA with two factors: condition (low vs. high imagery) and target ROI (LIPS vs. LT). This analysis revealed no main effect of condition or target ROI. However, the expected interaction between the factors was reliable [$F(1,10) =$

14.87, $P < 0.01$]. The intraparietal sulcus area had higher functional connectivities with the five ROIs for high-imagery sentences, whereas the left temporal area had higher functional connectivities to these five ROIs for low-imagery sentences, as shown in Fig. 6. This result provides an important converging operation for implicating the intraparietal sulcus in imagery processing in sentence comprehension, and more directly indicates the interaction of this processing with some of the other key activated regions.

Discussion

The findings substantiate that the left intraparietal sulcus area is associated with imagery in sentence comprehension. Additionally, two prefrontal areas, pars opercularis (bilaterally) and the left frontal eye fields (which are adjacent and superior to pars opercularis), also showed more activation in conjunction with the comprehension of high-imagery sentences. These areas may play a role in generating internal representations (perhaps including articulatory attributes) that are used in maintaining and communicating the sentence information. This maintenance and communication can be thought of as a working memory function that is used in sentence comprehension.

Two ROIs showed more activation in the low-imagery condition. In the left temporal ROI, the percentage of signal change was greater in the low-imagery condition. It may be that low-imagery sentences require more semantic processing of abstract concepts and this type of processing, associated with the left temporal area (Keller et al., 2001), engenders additional activation there. The

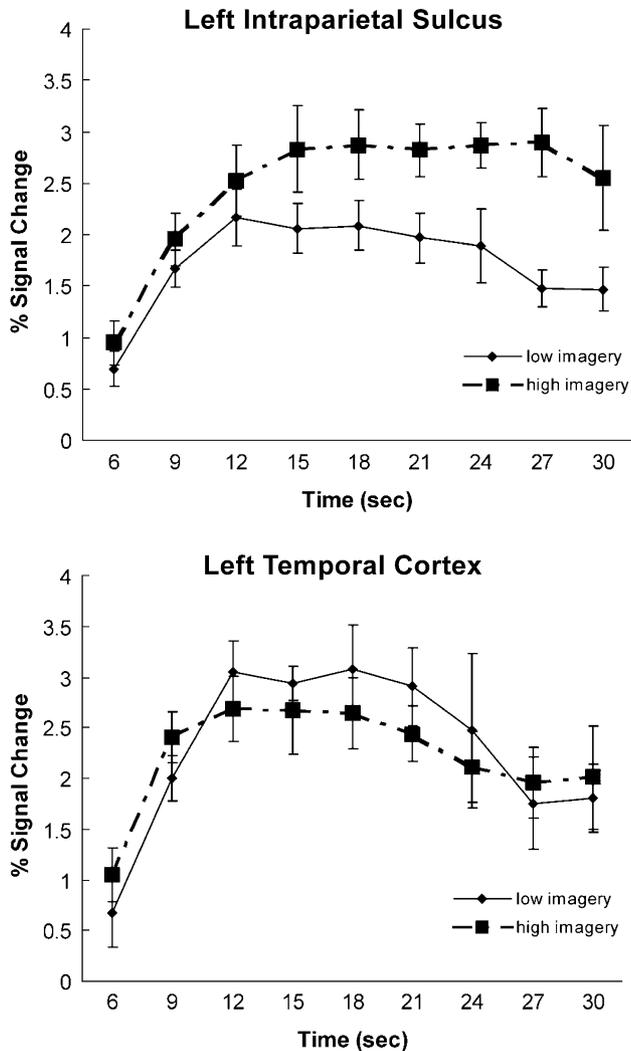


Fig. 5. Visual experiment time courses. The time courses from the activated voxels within left intraparietal sulcus and left temporal cortex, averaged across participants, with the standard error of the mean depicted.

other ROI with reliably more activation in the low-imagery condition (both in terms of the number of activated voxels and the signal intensity) was the calcarine sulcus. We speculate that the processing of the imagery content of high-imagery sentences was more consuming of resources within a visuospatial subnetwork of areas, thereby leaving fewer resources available for nonessential visual sensory processing, and hence less activation in the calcarine area. This speculation is consistent with the finding that the activation in V5 related to the perception of motion decreases in the presence of an additional cognitive processing demand (Rees et al., 1997).

The overall pattern of results raises interesting new questions about the brain activation effects of imagery during sentence comprehension. The new questions concern the possible effects of the visual encoding of written sentences. Might not some of the effects of the imagery manipulation (both positive and negative effects) be due to the interaction between imagery-related processes and the visuospatial processes (visual encoding and eye movements) involved in visual sentence comprehension? In fact, might not some of the activation in the IPS be due in part

to the spatial processing associated with the reading of the array of printed words, which could differ for the two types of sentences?

To examine this issue, a second study was run in which the same stimulus sentences were presented in the auditory modality, and the experimental conditions were otherwise as similar as possible.

Auditory experiment

Method

The participants listened to either high-imagery or low-imagery sentences to verify them. The answer options, “true” and “false,” were simultaneously displayed on a computer screen and the participant pressed one of two response buttons at the end of the sentence to indicate their decision. When they pressed a button, the next sentence was presented.

The sentences were the same as in the visual experiment, apart from two differences. First, because the auditory presentation approximately equated the presentation time for high- and low-imagery sentences, the number of sentences per high-imagery epoch was five, making it equal to the number of low-imagery sentences per epoch. Second, the response time was measured from the end of the sentence presentation.

Participants

The participants were 10 right-handed college students (three females) between 18 and 30 years old who gave signed informed consent approved by the University of Pittsburgh and Carnegie Mellon University Institutional Review Boards.

Results

Behavioral results

The accuracy rates were high and not reliably different between the high- and low-imagery sentences (88% and 90%, respectively; see Fig. 2). There was an imagery effect observed in the reaction

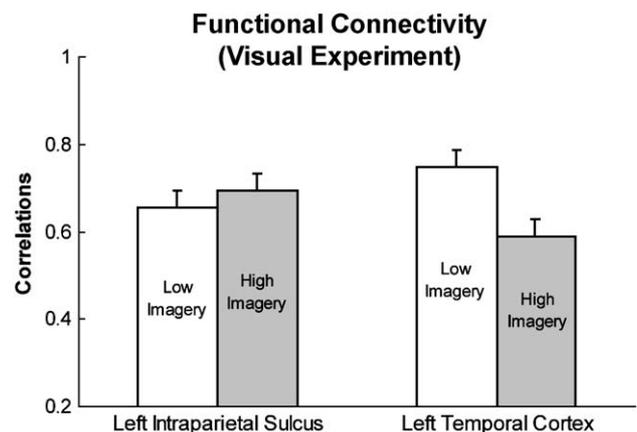


Fig. 6. Functional connectivity (visual experiment). The average functional connectivity of the left temporal ROI (LT) and the left intraparietal sulcus ROI (LIPS) with five left hemisphere ROIs: pars opercularis, pars triangularis, the inferior parietal lobule, dorsolateral prefrontal cortex, and frontal eye fields.

times, with the high-imagery sentences having a longer response time than the low-imagery sentences [means: 1353 vs. 786 ms; $F(1,9) = 17.65$, $P < 0.01$].

fMRI data overview

The auditory sentences experiment replicated the visual experiment in terms of the activation associated with the central imagery processing, while at the same time producing several activation differences associated with the differences in sensory–perceptual processing. The imagery effect observed in the intraparietal sulcus area during the processing of written sentences was clearly replicated in auditory sentence processing, with an even larger effect size than in visual sentence processing, as shown in Table 1. Fig. 7 shows the activation in two slices of one participant. Fig. 8 shows how the activation volumes were modulated by imagery in four key ROIs. Fig. 9 shows the time courses of the activation in the two contrasting ROIs, LIPS and LT. The frontal eye fields also continued to show the imagery effect despite the presentation being auditory. The imagery manipulation affected more ROIs in the listening condition: additional parietal and prefrontal and occipital ROIs were affected. As in the visual experiment, there were a few cortical regions with *less* activation in the high-imagery condition.

Spatial processing. In both the left and the right intraparietal sulcus ROIs, the high-imagery condition produced reliably more activation, both in terms of voxels activated and percentage of change in signal intensity. The effect size was 9 voxels in left

IPS, more than double the size in the reading experiment (4.3 voxels), as shown in Table 1. Thus, the IPS activation associated with imagery in comprehension occurs even more strongly in the absence of visual encoding processes (see Figs. 8 and 9). The imagery effect was more widespread in the parietal lobule in the listening compared to the reading condition. A significant imagery effect was observed not only in IPS, but also in the left superior parietal ROI, and in the left inferior parietal ROI.

As in the visual experiment, the frontal eye fields revealed significantly more activation in the high-imagery condition. However, the amount of activation and the imagery effect size was greater on the left (an increase from 5.4 to 10.1 voxels) than on the right (from 0.5 to 2.6 voxels). In the visual comprehension condition, the effect occurred only on the left, and the effect size was 2 voxels, much less than in the auditory condition. A possible interpretation is that the visuospatial processes in reading compete with imagery processing because they rely on similar neural substrates, suppressing the imagery effects. Imagery in the listening condition is uninhibited by such reading processes.

Language processing. Similar trends were observed in the pars opercularis during both the visual and the auditory experiments. However, the listening experiment failed to elicit a significant imagery effect in either left pars opercularis (although 7 of the 10 participants revealed more activation for the high-imagery condition) or right pars opercularis, as shown in Table 1 and Figs. 8 and 9. Significant effects of imagery were observed in

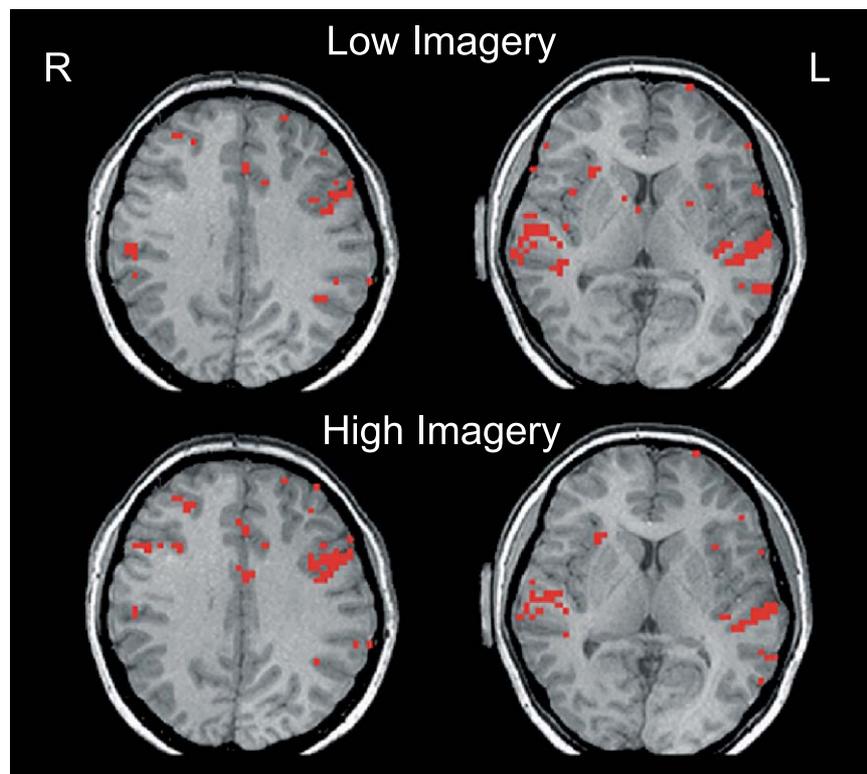


Fig. 7. Auditory experiment activation map. Activation from two axial slices (the right figure is more superior) of a typical participant (activation level approximately equal to group mean). The top two slices show activation resulting from the low-imagery condition and the bottom two the high-imagery condition.

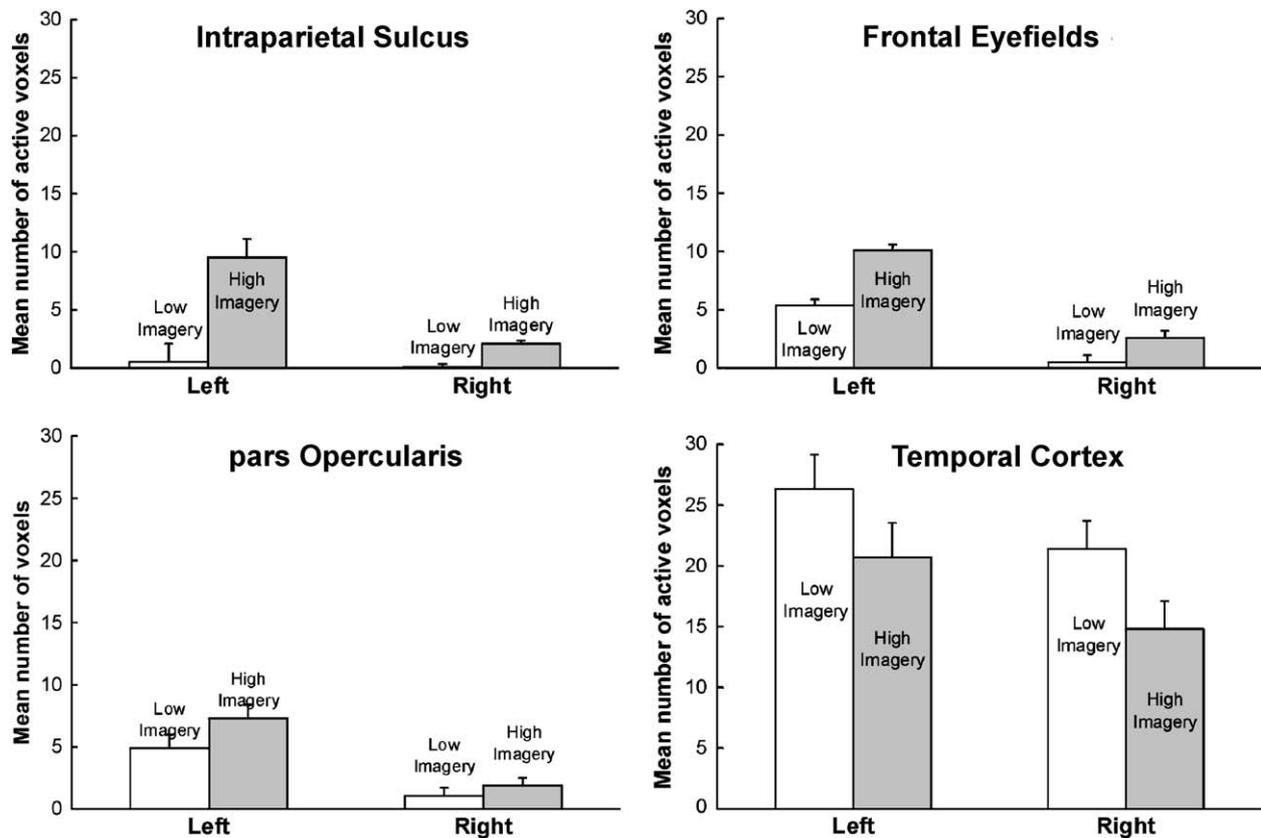


Fig. 8. Auditory experiment voxel counts. The mean number of activated voxels from the intraparietal sulcus, frontal eye fields, pars opercularis, and temporal cortex. Error bars represent 95% confidence intervals based on the pooled *Mse* from the corresponding ANOVA (Loftus and Mason, 1994).

the left DLPFC, which may be related to the increased working memory management and maintenance imposed by the listening condition.

As in the visual experiment, the temporal region revealed *more* activation during the low-imagery condition than during the high-imagery condition. This effect was observed in both the left and right temporal ROIs.

Primary sensory processing. The listening experiment, analogously to the visual one, revealed that the low-imagery sentences elicited *more* activation in primary auditory cortex than did the high-imagery sentences. Both left and right Heschl's gyrus (the primary auditory ROI) revealed an effect of imagery in the percentage of signal change measure. In addition, the auditory experiment also elicited activation within primary visual cortex. As in the visual experiment, there was *more* activation associated with the low-imagery sentences than the high-imagery sentences (as shown in Table 1).

The left extrastriate ROI revealed only a small amount of activation; nonetheless, there was a significantly reliable imagery effect, with more activation in high-imagery condition (see Table 1). In the visual experiment, the left extrastriate ROI revealed much more activation and a larger but unreliable imagery effect (from 12.6 to 15.3 voxels).

Laterality effects. Four cortical regions revealed reliably more activated voxels in the left hemisphere compared to the right. These regions included the: intraparietal sulcus [$F(1,9) = 20.02$, $P < 0.01$], frontal operculum [$F(1,9) = 14.28$, $P < 0.01$], pars

triangularis [$F(1,9) = 6.98$, $P < 0.01$], and frontal eye fields [$F(1,9) = 12.07$, $P < 0.01$]. Furthermore, there was a reliable interaction between hemisphere and imagery in the intraparietal sulcus [$F(1,9) = 11.03$, $P < 0.01$] and frontal eye fields [$F(1,9) = 7.21$, $P < 0.05$], indicating that high imagery was associated with reliably more activation in the left hemisphere ROI than in the right homologue. An interaction in the inferior parietal area was also observed [$F(1,9) = 6.23$, $P < 0.05$], indicating that in the left hemisphere high imagery was associated with more activation, whereas in the right hemisphere high imagery was associated with less activation.

Functional connectivity analysis. As in the visual experiment, the functional connectivity analysis revealed differential effects of sentence imagery on the functional connectivity of the left intraparietal sulcus and the left temporal cortex. The connectivities of the left intraparietal sulcus and the left temporal cortex with five left hemisphere ROIs (pars opercularis, pars triangularis, dorsolateral prefrontal cortex, frontal eye fields, and inferior parietal) were compared using a two-factor ANOVA: condition (low vs. high imagery) and target ROI (LIPS vs. LT). As in the visual experiment, this analysis revealed no significant main effects of condition or target ROI. However, the expected interaction was again reliable [$F(1,8) = 24.57$, $P < 0.01$]. The intraparietal sulcus area had higher functional connectivities to the five ROIs for high-imagery sentences, whereas the left temporal area had higher functional connectivities to these areas for low-imagery sentences, as shown in Fig. 10.

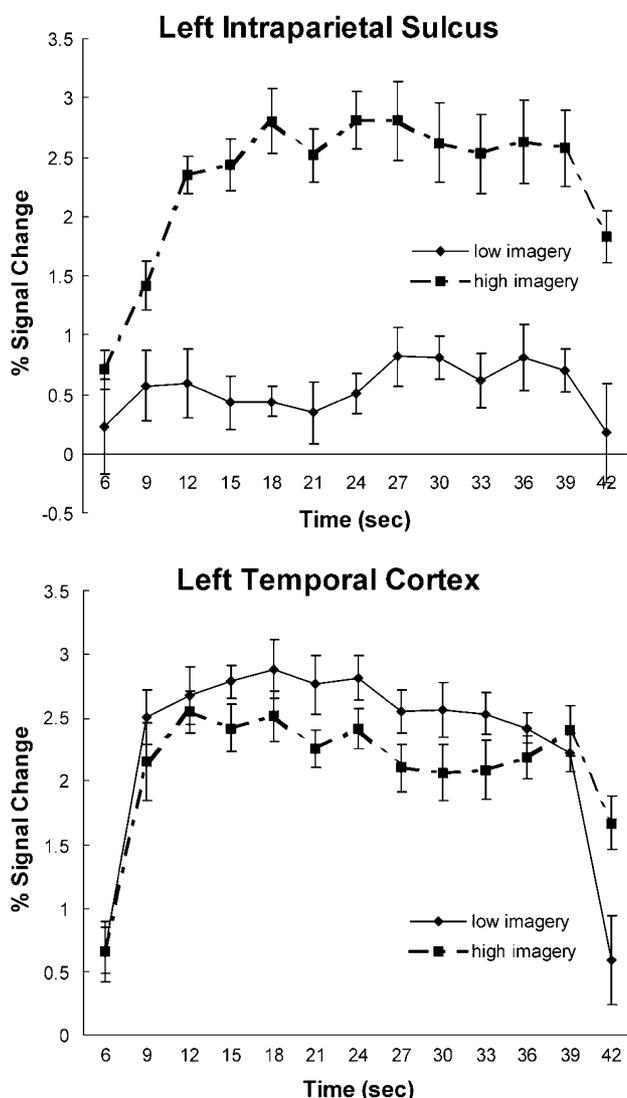


Fig. 9. Auditory experiment time courses. The time courses from the activated voxels within left intraparietal sulcus and left temporal cortex, averaged across participants, with the standard error of the mean depicted.

Modality differences. A between-subjects ANOVA compared the mean number of activated voxels in the visual and auditory experiments for each ROI, as well as assessing the imagery effects. As expected, the sensory areas, both calcarine and Heschl’s gyrus, showed a main effect of modality, but so did several association areas.

Several regions showed more activation during the visual experiment compared to the auditory experiment. They included the right intraparietal sulcus [$F(1,20) = 11.31, P < 0.01$], left and right extrastriate cortex [left: $F(1,20) = 18.51, P < 0.01$; right: $F(1,20) = 9.73, P < 0.01$], and the calcarine sulcus [$F(1,20) = 25.29, P < 0.01$].

Other regions were found to be more active during the auditory experiment. They included the right inferior parietal region [$F(1,20) = 5.43, P < 0.05$], the left frontal eye fields [$F(1,20) = 10.95, P < 0.01$], the left pars triangularis [$F(1,20) =$

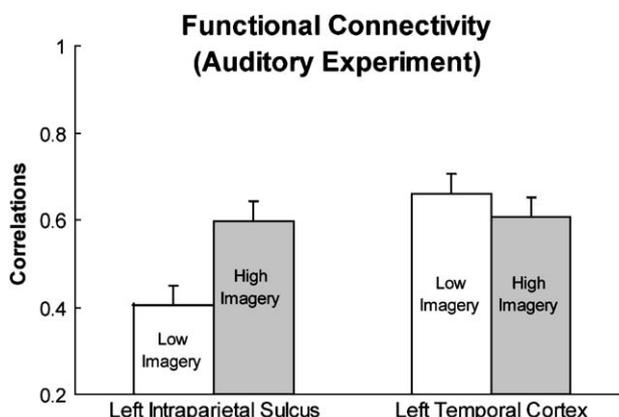


Fig. 10. Functional connectivity (auditory experiment). The average functional connectivity of the left temporal ROI (LT) and the left intraparietal sulcus ROI (LIPS) with five left hemisphere ROIs: pars opercularis, pars triangularis, the inferior parietal lobule, dorsolateral prefrontal cortex and frontal eye fields.

4.5, $P < 0.05$], the left and right temporal regions [left: $F(1,20) = 11.23, P < 0.01$; right: $F(1,20) = 27.08, P < 0.01$], and left and right Heschl’s gyrus [left: $F(1,20) = 32.35, P < 0.01$; $F(1,20) = 49.15, P < 0.01$].

Of these regions, the left frontal eye fields [$F(1,20) = 8.28, P < 0.01$], right superior parietal cortex [$F(1,20) = 4.75, P < 0.05$], the right temporal region [$F(1,20) = 4.82, P < 0.05$], and the calcarine sulcus [$F(1,20) = 10.03, P < 0.01$] also showed significant imagery by modality interactions while the left intraparietal sulcus revealed a marginally significant interaction [$F(1,20) = 3.36, P = 0.08$].

Laterality and activation centroids differed as a function of modality, as shown in Tables 1 and 2. Laterality differences were found in the regions involved in spatial processing (revealing greater left lateralization during the auditory experiment) and in the temporal cortex (revealing greater left lateralization during the visual experiment). The activation centroids differed in four ROIs which would be expected to be affected by the nature of the sensory input, namely, the left and right temporal cortex (which are secondary auditory processing areas) and the left and right extrastriate areas (which are secondary visual processing areas). In addition, the activation locations differed between the auditory and visual experiments in both left and right DLPFC. While previous studies have indicated location differences within DLPFC as a function of processing domain (e.g., verbal, spatial, or object) differences, this is one of the first experiments to report differences in this region due to presentation modality.

Table 2
Laterality ratios (left – right/left + right) for selected ROIs

ROI	Visual	Auditory
Intraparietal sulcus	0.27	0.65
Frontal eye fields	0.30	0.71
Pars opercularis	0.62	0.61
Temporal cortex	0.88	0.13

General discussion

The results of both the auditory and visual comprehension experiments showed that certain cortical regions involved in spatial processing (particularly the left intraparietal sulcus) were activated to a greater degree during the processing of high-compared to low-imagery sentences. Moreover, the left intraparietal sulcus showed greater functional connectivity with other regions during the processing of high-imagery sentences. It has been hypothesized that the left parietal region may be a visuo-spatial workspace and that some aspects of mental imagery are lateralized to the left hemisphere (D'Esposito et al., 1997; Farah et al., 1988; Ishai et al., 2000). For example, in a study in which participants were to either view images of objects or generate mental images of objects, imagery elicited significantly more left parietal activation, particularly in and around the intraparietal sulcus, than in its right homologue (Ishai et al., 2000). The role of the intraparietal sulcus might be to construct the specified spatial representations. Activation in this area is often observed when a spatial representation is generated, manipulated, or interrogated (Baker et al., 1996; Carpenter et al., 1999; Dagher et al., 1999; Diwadkar et al., 2000; Just et al., 2001; Nystrom et al., 2000).

In addition to the left intraparietal sulcus region, the left frontal eye fields and left pars opercularis showed more activation during the comprehension of high-imagery sentences compared to low-imagery sentences, reminiscent of Mellet et al. (1998). In that study, the high-imagery condition elicited greater involvement of the region in and around the intersection of the precentral gyrus and the inferior frontal gyrus, which corresponds approximately to the ROI called the frontal eye fields in our parcellation scheme. This area has been thought to be involved in eye movement control, especially saccadic eye movements (Colby, 1991) and the control of spatial attention (particularly the right frontal eye fields) (Coull and Frith, 1998), but the higher cognitive functions of this area, particularly the functions related to language processing, have not been specified. In addition to showing imagery effects, this region also revealed a leftward asymmetry, which has been observed previously in both visual and auditory sentence comprehension tasks (Keller et al., 2001; Michael et al., 2001).

We suggest that in the collaboration among the cortical regions that process imagery, the region at the intersection of the inferior frontal sulcus and the precentral sulcus (left pars opercularis and frontal eye fields) might bear responsibility for generating an internal nonspatial data structure that contains the information to be “plotted” in the spatial representation. This nonspatial data structure could well contain pointers to stored spatial representations that do contain spatial information, such as the specification of the shape of the numeral 8. This hypothesis is consistent with other researchers' conceptions of the function of this region, which can be construed broadly as “structure building.” For example, in a study in which participants imagined movement trajectories, activation of the left pars opercularis was observed (Binkofski et al., 1997, 2000). In that study, structure building could include the generation of an internal data structure that represents the trajectory at a particular orientation and the intermediate representations of that trajectory. In processing a high-imagery sentence, on the other hand, these regions may be involved in the processing of the sequence of words which, when syntactically and semantically interpreted, provide the contents

for the data structure that defines the image. The mental data structure attributed to the left frontal region is proposed to be communicated to the left intraparietal sulcus area, where a spatial representation is constructed.

Furthermore, we suggest that the processes we attribute to this collaboration between the frontal (left pars opercularis–frontal eye fields) and parietal (left intraparietal sulcus) regions occur iteratively and interactively, rather than in isolation from each other. For example, after the prefrontal region specifies some aspect of the spatial representation, this information might be communicated to the intraparietal sulcus area, and there may be subsequent feedback to pars opercularis to further elaborate some aspect of the specification. Thus, there could be a verbal–spatial network that mediates between the construction of the two types of representations that are often both used in language comprehension. The functional connectivity results (discussed below) provide converging evidence for this proposal.

While visuospatial processing regions revealed greater involvement during the comprehension of high-imagery sentences, semantic processing regions (i.e., superior and middle temporal cortex) revealed greater activation during the comprehension of low-imagery sentences. A previous neuroimaging study examining the interactions between language and visual imagery revealed somewhat similar results (Mellet et al., 1998). In that PET study, the high-imagery condition consisted of listening to concrete word definitions and generating their mental image while the low-imagery condition required listening to abstract word definitions. As in the current study, Mellet et al. also observed greater involvement of temporal cortex during the low-imagery condition. The posterior middle and superior temporal cortex is believed to subservise interpretive and elaborative functions (Keller et al., 2001). One possible explanation for the increased activation of temporal cortex during the comprehension of low-imagery sentences is that low-imagery sentences require more lexical–semantic processing to determine their truth value than do high-imagery sentences, which require more spatial processing.

Functional connectivity

The functional connectivity findings provide substantial converging support for the role of the intraparietal sulcus in the processing underlying sentence imagery. Not only the activation level but also the degree of synchronization of the IPS with other activated regions (including those involved in comprehension) is very much a function of whether imagery processing is required. This finding was replicated in the visual and auditory experiments.

A related functional connectivity finding is that the left temporal area showed the opposite pattern of connectivities to LIPS, indicating the greater involvement of the left temporal area in the processing of low-imagery sentences. Moreover, the modality of presentation appears to influence the functional connectivity patterns. There is a greater difference in functional connectivity between the two imagery conditions in LIPS in the auditory experiment, whereas there is a greater difference in functional connectivity in LT in the visual experiment. A tentative implication of this pattern is that listening comprehension may be a better modality in which to receive imagery-laden information, whereas the reading comprehension may be a better modality in which to receive abstract information.

Modality effects

The juxtaposition of the visual and auditory experiments allows us to compare the neural bases of the conceptual (nonsensory) aspects of sentence comprehension in reading vs. listening comprehension. The transient nature of the auditory stimulus in listening comprehension inherently imposes greater demands on verbal working memory and therefore may evoke more semantic processing to accommodate the greater memory demands of listening. Both of these functions (semantic processing and verbal working memory) have been associated with the left pars triangularis (Fiez, 1997; Fiez and Petersen, 1998; Gabrieli et al., 1998; Petersen et al., 1989, 1990). Consistent with this view, the current experiments found more activation in left pars triangularis during listening comprehension than reading comprehension, as was previously found by Michael et al. (2001).

In summary, this first attempt at examining the neural bases of the interaction between complex language comprehension and mental imagery provides new insights into how two apparently separable processes function in coordination. The results indicate that high- and low-imagery sentences evoke a common set of cortical areas, but differentially rely on these areas. High-imagery sentences evoke more visuospatial processing in the left intraparietal sulcus area, whereas low-imagery sentences evoke more semantics-related processing in the left temporal area. Of equal significance is the higher functional connectivity of these two areas with other key areas when their specialization is called into play.

Acknowledgments

This research was supported in part by the Office of Naval Research Grant N00014-96-1-0322 and the National Institute of Mental Health Research Scientist Award MH-00662.

Appendix A. Stimuli

High-imagery sentences

1. The letter W rotated 180 degrees and written in lower case looks like the letter m.
2. The number eight when rotated 90 degrees looks like a pair of spectacles.
3. The capital letter a is formed with two diagonal lines and a vertical line.
4. The letter c attached to the left side of a square resembles the outline of a cup.
5. The number nine can be constructed from a circle and a horizontal line.
6. On a map, south is usually the compass direction that points toward the top.
7. A circle placed on top of the capital letter v resembles the outline of an ice cream cone.
8. The letter V with a vertical line placed on top is an arrow pointing downward.
9. The hands on a clock at 9 o'clock look like the capital letter L.
10. In terms of diameter, a quarter is larger than a nickel, which is larger than a dime.
11. A capital letter n rotated 90 degrees clockwise, results in the capital letter z.
12. On a map, Nevada is to the right of California.

13. If one rotates the small letter B by 180 degrees, the result is the letter P.
14. On a map of Europe, Spain is to the left of France and France is below Italy.
15. A square with the base of a triangle on top is the outline of a traditional house.

Low-imagery sentences

1. Horsepower is the unit for measuring the power of engines or motors.
2. To the citizens of ancient Greece, Vishna was considered the supreme deity.
3. Economics deals with the production, distribution, and consumption of wealth.
4. Baptism in the Christian faith is considered a means of washing away original sin.
5. Thomas Edison was an American scientist who invented the first computer.
6. A galaxy is a grouping of millions of stars merging into a luminous band.
7. Italian, French, and Japanese are considered to be European romance languages.
8. More like a violin than a tuba, the piano is actually a stringed instrument.
9. Human beings have keener powers of vision than birds such as hawks and eagles.
10. The gestation of a human fetus takes approximately 9 months to complete.
11. In California, raisins are made by drying out the white and red table grapes.
12. Although now a sport, marathons started with Greek messengers bringing news.
13. It is not dangerous if a woman contracts German measles during early pregnancy.
14. Mammals, including dogs and cats, are vertebrates that hatch their young in eggs.
15. The first plays that we know of that were performed in Europe were of Greek origin.

References

- Baker, S., Rogers, R.D., Owen, A.M., Frith, C.D., Dolan, R.J., Frackowiak, R.S.J., Robbins, T.W., 1996. Neural systems engaged by planning: a PET study of the Tower of London task. *Neuropsychologia* 34, 515–526.
- Binkofski, F., Stephan, K.M., Posse, S., Shah, N.J., Muller-Gartner, H.W., Freund, H.J., Seitz, R.J., 1997. Imagery strategies in kinematic processing determine cortical activations. *NeuroImage* 54, 127 (Abstract).
- Binkofski, F., Amunts, K., Stephan, C.M., Posse, S., Schormann, T., Freund, H., Zilles, K., Seitz, R.J., 2000. Broca's region subserves imagery of motion: a combined cytoarchitectonic and fMRI study. *Hum. Brain Mapp.* 11, 273–285.
- Buchel, C., Coull, J.T., Friston, K.J., 1999. The predictive value of changes in effective connectivity for human learning. *Science* 283, 1538–1541.
- Carpenter, P.A., Just, M.A., Keller, T.A., Eddy, W., Thulborn, K., 1999. Graded functional activation in the visuospatial system with the amount of task demand. *J. Cogn. Neurosci.* 11, 9–24.
- Caviness Jr., V.S., Meyer, J., Makris, N., Kennedy, D.N., 1996. MRI-based topographic parcellation of human neocortex: an anatomically specified method with estimate of reliability. *J. Cogn. Neurosci.* 8, 566.
- Colby, C.L., 1991. The neuroanatomy and neurophysiology of attention. *J. Child Neurol.* 6, 90–118.

- Coull, J.T., Frith, C.D., 1998. Differential activation of right superior parietal cortex and intraparietal sulcus by spatial and nonspatial attention. *NeuroImage* 8, 176–187.
- Dagher, A., Owen, A., Boecker, H., Brooks, D., 1999. Mapping the network for planning: a correlational PET activation study with the Tower of London. *Brain* 122, 1973.
- D'Esposito, M., Detre, J.A., Aguirre, G.K., Stallcup, M., Alsop, D.C., Tippet, L.J., Farah, M.J., 1997. A functional MRI study of mental image generation. *Neuropsychologia* 35, 725–730.
- Dieber, M.-P., Ibanez, V., Honda, M., Sadato, N., Raman, R., Hallett, M., 1998. Cerebral processes related to visuomotor imagery and generation of simple finger movements studies with positron emission tomography. *NeuroImage* 7, 73–85.
- Diwadkar, V.A., Carpenter, P.A., Just, M.A., 2000. Collaborative activity between parietal and dorso-lateral prefrontal cortex in dynamic spatial working memory revealed by fMRI. *NeuroImage* 12, 85–99.
- Eddy, J.K., Glass, A.L., 1981. Reading and listening to high and low imagery sentences. *J. Verbal Learn. Verbal Behav.* 20, 333–345.
- Eddy, W., Fitzgerald, M., Genovese, C., Mockus, A., Noll, D., 1996. Functional imaging analysis software—Computational olo. *Proceedings in Computational Statistics*. Physica-Verlag, Heidelberg, pp. 39–49.
- Farah, M.J., Hammond, D.M., Levine, D.N., Calvanio, R., 1988. Visual and spatial mental imagery: dissociable systems of representation. *Cogn. Psychol.* 20, 439–462.
- Fiez, J.A., 1997. Phonology, semantics, and the role of the left inferior prefrontal cortex. *Hum. Brain Mapp.* 5, 79–83.
- Fiez, J.A., Petersen, S.E., 1998. Neuroimaging studies of word reading. *Proc. Natl. Acad. Sci. U.S.A.* 95, 914–921.
- Friston, K.J., 1994. Functional and effective connectivity in neuroimaging: A synthesis. *Hum. Brain Mapp.* 2, 56–78.
- Gabrieli, H.D.E., Poldrack, R.A., Desmond, J.E., 1998. The role of the left prefrontal cortex in language and memory. *Proc. Natl. Acad. Sci. U. S. A.* 95, 906–913.
- Horwitz, B., Rumsey, J.M., Donohue, B.C., 1998. Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proc. Natl. Acad. Sci. USA* 95, 8939–8944.
- Ishai, A., Ungerleider, L.G., Haxby, J.V., 2000. Distributed neural systems for the generation of visual images. *Neuron* 28, 979–990.
- Just, M.A., Carpenter, P.A., Maguire, M., Diwadkar, V., McMains, S., 2001. Mental rotation of objects retrieved from memory: an fMRI study of spatial processing. *J. Exp. Psychol. Gen.* 130, 493–504.
- Keller, T.A., Carpenter, P.A., Just, M.A., 2001. The neural bases of sentence comprehension: an fMRI examination of syntactic and lexical processing. *Cereb. Cortex* 11, 223–237.
- Kosslyn, S.M., Alpert, N.M., Thompson, W.L., Maljkovic, V., Weise, S.B., Chabris, C.F., Hamilton, S.E., Rauch, S.L., Buonanno, F.S., 1993. Visual mental imagery activates topographically organized visual cortex: PET investigations. *J. Cogn. Neurosci.* 5, 263–287.
- Kosslyn, S.M., Shin, L.M., Thompson, W.L., McNally, R.J., Rauch, S.L., Pitman, R.K., Alpert, N.M., 1996. Neural effects of visualizing and perceiving aversive stimuli: a PET investigation. *NeuroReport* 7, 1569–1576.
- Kosslyn, S.M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J.P., Thompson, W.L., Ganis, G., Sukel, K.E., Alpert, N.M., 1999. The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167–170.
- Lazar, N.A., Eddy, W.F., Genovese, C.R., Welling, J.S., 2001. Statistical issues in fMRI for brain imaging. *Int. Stat. Rev.* 69, 105–127.
- Loftus, G.R., Mason, M.E.J., 1994. Using confidence intervals in within-subject designs. *Psychol. Bull. Rev.* 1, 1476–1490.
- Mazoyer, B., Tzourio-Mazoyer, N., Mazard, A., Denis, M., Mellet, E., 2002. Neural bases of image and language interactions. *Int. J. Psychol.* 37, 204–208.
- Mellet, E., Tzourio, N., Crivello, F., Joliot, M., Denis, M., Mazoyer, B., 1996. Functional neuroanatomy of spatial mental imagery generated from verbal instructions. *J. Neurosci.* 16, 6504–6512.
- Mellet, E., Tzourio, N., Denis, M., Mazoyer, B., 1998. Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition. *NeuroReport* 9, 803–808.
- Mellet, E., Tzourio-Mazoyer, N., Bricogne, S., Mazoyer, B., Kosslyn, S.M., Denis, M., 2000. Functional anatomy of high-resolution visual mental imagery. *J. Cogn. Neurosci.* 12, 98–109.
- Mellet, E., Bricogne, S., Crivello, F., Mazoyer, B., Denis, M., Tzourio-Mazoyer, N., 2002. Neural basis of mental scanning of a topographic representation built from a text. *Cereb. Cortex* 12, 1322–1330.
- Michael, E.B., Keller, T.A., Carpenter, P.A., Just, M.A., 2001. An fMRI investigation of sentence comprehension by eye and by ear: modality fingerprints on cognitive processes. *Hum. Brain Mapp.* 13, 239–252.
- Nystrom, L.E., Braver, T.S., Sabb, F.W., Delgado, M.R., Noll, D.C., Cohen, J.D., 2000. Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. *NeuroImage* 11, 424–446.
- Petersen, S.E., Fox, P.T., Posner, M.I., Mintun, M., Raichle, M.E., 1989. Positron emission tomographic studies of the processing of single words. *J. Cogn. Neurosci.* 1, 153–170.
- Petersen, S.E., Fox, P.T., Snyder, A.Z., Raichle, M.E., 1990. Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science* 249, 1041–1044.
- Rademacher, J., Galaburda, A.M., Kennedy, D.N., Flipepek, P.A., Caviness, V.S., 1992. Human cerebral cortex: localization, parcellation, and morphometry with magnetic resonance imaging. *J. Cogn. Neurosci.* 4, 352.
- Rees, G., Frith, C.D., Lavie, N., 1997. Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science* 278, 1616–1619.