Website publication 11 March 1998

NeuroReport 9, 803–808 (1998)

Cortical anatomy of mental imagery of concrete nouns based on their dictionary definition

E. Mellet, N. Tzourio, M. Denis[1] and B. Mazoyer[CA]

UPRES 2127 Université de Caen and CEA LRC 13, GIN GIP Cyceron, BP 5229, 14074 Caen Cedex; [1]LIMSI, CNRS, Orsay, France

[CA]Corresponding Author

THE functional anatomy of the interactions between spoken language and visual mental imagery was investigated with PET in eight normal volunteers during a series of three conditions: listening to concrete word definitions and generating their mental images (CONC), listening to abstract word definitions (ABST) and silent REST. The CONC task specifically elicited activations of the bilateral inferior temporal gyri, of the left premotor and left prefrontal regions, while activations in the bilateral superior temporal gyri were smaller than during the ABST task, during which an additional activation of the anterior part of the right middle temporal gyrus was observed. No activation of the occipital areas was observed during the CONC task when compared either to the REST or to the ABST task. The present study demonstrates that a network including part of the bilateral ventral stream and the frontal working memory areas is recruited when mental imagery of concrete words is performed on the basis of continuous spoken language *NeuroReport* 9: 803–808 © 1998 Rapid Science Ltd.

Key words: Fusiform gyrus; Inferior temporal cortex; Language comprehension; Mental imagery; PET; Working memory; Prefrontal cortex

Introduction

The human brain can rely on two main systems for mental representation: language and visual mental imagery.[1] While language conveys symbolic representations allowing human communication, mental imagery operates on internal representations and as such is a major component of human thought. Strong interactions exist between these two systems that permit an easy transfer from one representation mode to the other: one can indeed easily build a mental image from a verbal description[2,3] or, conversely, give an oral description of a mental image.

The neural bases of visual mental imagery have been extensively studied over the past several years using positron emission tomography (PET)[3–7] and functional magnetic resonance imaging (fMRI).[8,9] However, sparse and conflicting results have been reported on the functional anatomy of the interactions between language and mental imagery: some studies have found an activation of the primary and/or associative visual cortex when the mental image was generated from either a single word[7,9] or verbal instructions,[3] whereas others have found a deactivation of the same areas when mental evocation was guided by verbal instructions.[10,11] It is noticeable, however, that in all these studies, the language stimulus was restricted to single words or simple verbal instructions, making it quite different from usual speech. In addition, all these studies were designed such that listening to the language stimulus and mental image generation were subsequent to one another rather than simultaneous. The goal of the present study was thus to investigate the cortical areas involved during mental imagery based on simultaneous listening of continuous speech stimuli: regional cerebral blood flow was monitored by PET in subjects who were asked to generate and modify a visual mental image of a concrete word while listening to its definition, a task mixing to a high degree language listening and comprehension and mental imagery.

Subjects and Methods

Subjects: Eight right-handed young male healthy volunteers (aged 20–25 years) participated in the study. They were selected as 'high visuo-spatial imagers' according to the score they obtained on the Mental Rotations Test[12] (MRT; mean score (± s.d.) 16.5 ± 2.2) and the Minnesota Paper Form Board[13] (MPFB; mean score 22.5 ± 3.6), corresponding for both tests to the upper 33rd percentile of 106 male reference subjects. Each subject had a normal magnetic resonance imaging (MRI) brain scan and gave his written informed consent to the PET protocol that was accepted by the local ethical committee.

Tasks design: Using our standard ^{15}O-labeled water PET activation protocol,[3] nine sequential measurements of the normalized cerebral blood flow (NrCBF) were obtained on an ECAT Exact HR+ PET camera from each subject, replicating three times a series of three experimental conditions presented in a randomized order: (1) listening to the definition of concrete word and generating the corresponding mental image (CONC); (2) listening to the definition of abstract words (ABST); (3) a rest condition (REST).

In both the CONC and the ABST tasks the subjects were instructed to attentively listen to and understand 15 words and their definition, taken from a French dictionary (Larousse, 1993), verbally delivered through earphones. Each word/definition couple enunciation lasted 6 s and was followed by a 2 s silence before the next couple was delivered. The task duration was 120 s, starting 30 s before and maintained during the 90 s of the PET data acquisition. Different words were used for each list, the total number of words including those constituting the definitions being not different across the six definition lists ($p = 0.83$, ANOVA).

The words delivered during the CONC condition were in common use and easy to image, referring to objects or animals (e.g. bottle, guitar, lion). The characteristic of their definition was that it described the figural, the physical, or the functional features of the object or of the animal the word referred to. The definitions were thus very likely to result in spontaneous visual imagery activity. In addition, in order to induce sustained mental imagery, the subjects were explicitly encouraged to produce visual images evoked by the word and to modify or render more accurate this image as they listened to the definition following the word enunciation.

In the ABST condition, the words constituting the lists were part of the usual vocabulary referring to abstract notions. Their definitions also used usual abstract words and corresponded to concepts that were very unlikely to result in visual mental image spontaneous generation (e.g. grammar, theory, synthesis). Moreover, the subjects were instructed not to force themselves to produce mental images.

The REST control condition consisted of lying silently with eyes closed, as in all conditions, with no particular instruction except to refrain from moving.

Post-PET control of the tasks: Although the subjects received no memorization instruction before either task, they were asked immediately after each CONC or ABST task completion, to recall as many words as they could among the list of 15 they had just listened to. The word recall scores were collected as an indirect control of the visual mental imagery during the CONC condition since, according to the literature,[1] performances in the retrieval of words are better for concrete than for abstract words, thanks to the greater mental imageability of the former. During the same post-PET session, the subjects understanding of the definitions was evaluated by having them retrieve three randomly chosen words of the list from their definitions.

Data analysis: Global differences in rCBF were removed by scaling and statistical parametric maps (SPM) corresponding to comparisons between the CONC, ABST and REST conditions were generated using the SPM 96 software package.[14] A three-level task factor ANOVA was first performed, the corresponding F-map being thresholded at $p = 0.001$ uncorrected for multiple comparisons. *Post-hoc* t-statistic maps were then generated for the CONC minus REST, the CONC minus ABST and the ABST minus CONC contrasts and the corresponding Z volumes thresholded at $Z = 2.33$ ($p = 0.01$).

Results

Behavioral results: During the post PET free recall control, the subject significantly better retrieved concrete words (9.5 ± 2.3) than abstract words (6.6 ± 2.2; $p = 0.002$, paired t-test). The number of retrieved words was not different between the three concrete definition lists ($p = 0.38$, one-factor ANOVA), or between the three abstract definitions lists ($p = 0.53$).

Performance on the definition cued recall of words was perfect for concrete words (3.0 ± 0.0) and slightly but significantly lower for abstract words (2.5 ± 0.6, $p = 0.008$, paired t-test).

PET results: The stereotactic coordinates and the spatial extent of the CONC minus REST contrast are shown in Table 1, and in Table 2 for both the CONC minus ABST and the ABST minus CONC comparisons. The three corresponding Z volumes are displayed in Fig. 1. Compared with REST, the CONC task elicited bilateral activations of the whole superior temporal gyri, including the temporal poles and the Heschl's gyri, the posterior limit of the superior temporal activation being located vertically at the Sylvian fissure. A significant bilateral activation, albeit larger in the left hemisphere, was detected in a region that included the posterior part of the inferior temporal gyrus and the posterior part of the fusiform gyrus. A significant distinct activation was also detected in the left fusiform gyrus, in the occipital

Table 1. Foci of significant NrCBF increases in the CONC condition compared with the REST condition.

Region size	Anatomical location of max. voxel	Coordinates			Z score
		x	y	z	
Concrete definitions minus Rest					
3533	L. superior temporal sulcus	–56	–30	6	7.5
	L. superior temporal gyrus	–64	–8	4	7.0
	L. temporal pole	–58	10	–6	5.9
2719	R. superior temporal gyrus	66	–14	6	7.5
	R. superior temporal sulcus	64	–4	0	7.3
	R. superior temporal sulcus	60	–28	4	7.1
90	L. premotor	–38	0	64	4.9
39	SMA	12	4	48	4.4
72	L. precentral/middle frontal sulcus	–36	6	32	4.1
	L. precentral/middle frontal sulcus	–28	12	30	3.8
30	R. inferior temporal/fusiform gyrus	44	–50	–12	3.8
342	L. inferior temporal/fusiform gyrus	–48	–60	–14	3.6
	L. inferior temporal/fusiform gyrus	–56	–62	–14	3.5
39	L. fusiform gyrus	–44	–24	–20	3.2
28	R. anterior insula	44	18	2	28

The data, based on eight subjects, are local maxima detected with SPM software. Activated region volumes are given in voxels. Within these regions, the anatomical localization of the maximum Z scores of the voxel is given on the basis of the Talairach and Tournoux atlas, using their stereotactic coordinates in mm (R., right; L., left).

Table 2. Foci of significant NrCBF increases in the CONC condition compared with the ABST condition and significant NrCBF increase in the ABST condition compared with the CONC condition (for details, see Table 1 footnote).

Region size	Anatomical location of max. voxel	Coordinates			Z score
		x	y	z	
Concrete definition minus Abstract definition					
131	L. inferior temporal/fusiform gyrus	–42	–32	–18	4.6
342	L. inferior temporal gyrus	–52	–62	–6	4.3
	L. fusiform gyrus	–44	–58	–22	4.0
87	R. inferior temporal/fusiform gyrus	52	–50	–14	4.2
72	L. precentral/middle frontal sulcus	–40	4	34	4.0
	L. precentral/middle frontal sulcus	–28	14	30	3.2
45	L. inferior parietal lobule	–46	–38	46	4.0
43	L. precentral gyrus	–42	–16	38	3.3
55	L. premotor	–36	4	60	3.2
Abstract definition minus Concrete definition					
77	R. middle temporal gyrus	58	2	–18	3.9
292	R. superior temporal sulcus	54	–20	6	3.4
	R. superior temporal gyrus	52	–26	18	3.1
109	L. superior temporal gyrus	–60	–22	12	3.1

lobe. The left temporal lobe activated volume extended anteriorily to the inferior frontal gyrus, corresponding to Broca's area. In the frontal lobe, the CONC minus REST contrast also revealed a significant activation in the left lateral premotor cortex, and of the right pre-SMA. A cluster of activated voxels was also detected at the intersection between the left middle frontal sulcus and the precentral sulcus. Finally, the anterior part of the right insula was significantly activated in this contrast.

Comparison between the CONC and the ABST conditions extinguished the superior temporal gyri activations, while the bilateral activations of the inferior temporal and fusiform gyri persisted. In the left hemisphere, the volume of activated area extended to the vicinity of the middle occipital gyrus. The left prefrontal region lying within the depth of the middle frontal sulcus also remained activated. Moreover this comparison revealed activations in the left precentral gyrus and in the inferior parietal lobule but that was due to larger deactivations during the ABST than during the CONC condition when compared to REST.

The ABST minus CONC comparison showed bilateral activations of the superior temporal gyri with a rightward asymmetry. Moreover, in the

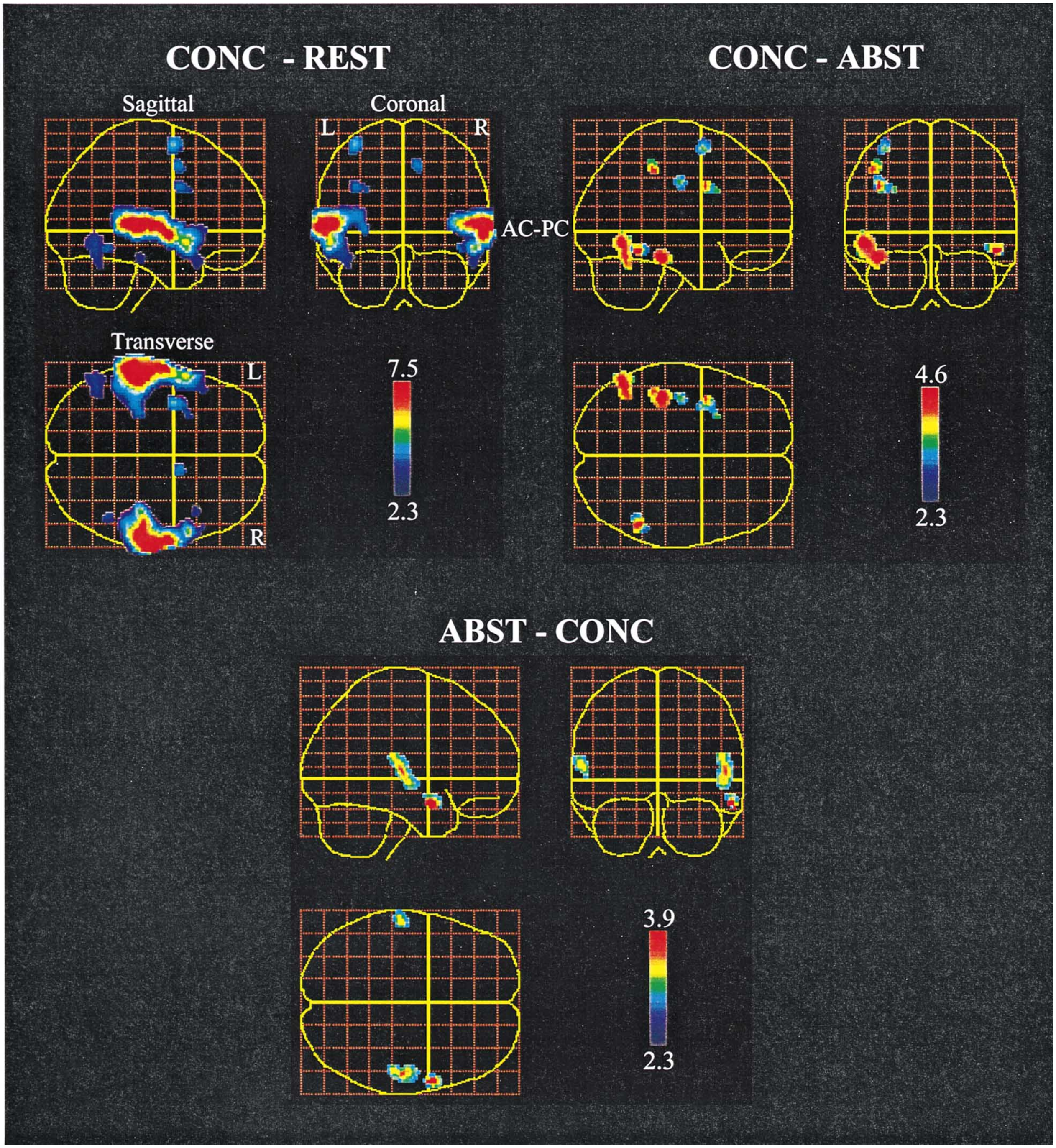

FIG. 1. SPMs corresponding to the CONC minus REST condition comparison (left), to the CONC minus ABST condition comparison (right) and to the ABST minus CONC condition comparison (bottom). Z volumes are projected in three orthogonal directions.

anterior part of the right middle temporal gyrus, along the superior temporal sulcus, a significant activation was detected due to a greater NrCBF increase during ABST than during CONC.

Discussion

The higher performance on the recall of concrete words compared with recall of abstract words indicates that the CONC condition was accompanied by a sustained visual imagery activity compared with the ABST condition.[1] Furthermore, the very good performances in the post-PET recall task of a word by its definition show indirectly that the subjects paid a sustained attention and actually understood the definitions during both the CONC and the ABST tasks. The CONC condition can therefore be assumed to correspond to a cognitive task during which listening and understanding spoken language and visual imagery were closely intermingled.

Bilateral activation of the inferior temporal and the fusiform gyri: Bilateral activation at the confluence of the inferior temporal gyrus and of the fusiform gyrus was detected both in CONC minus REST and CONC minus ABST contrasts, showing that this activation was specific of the listening condition with visual imagery. This cortical region belongs to the ventral pathway known to be involved in the identification of objects and faces,[15,16] and more generally seems to play a part in storing and evoking the figurative aspects of visual representations.[17] Actually, the inferior temporal and adjacent fusiform gyri have been found to be activated in most PET and fMRI studies on generation of visual mental images.[3–5,9]

While in agreement with most of these previous findings, the bilateral inferior temporal cortex activation observed in our study questions the existence of a leftward functional lateralization of the inferior temporal cortex for the generation of a mental image, as has been proposed by others.[9] In the visual perception domain, the right temporal gyrus and the adjacent fusiform gyrus are activated during the identification of complex shapes.[18,19] This characteristic has also been observed in visual imagery studies: in a previous study we reported an activation of this region in the right hemisphere during the mental construction of complex shapes.[3] The visualization of a letter on a grid also generated a right-sided activation,[4] as was also the case with the visualization of objects limiting a mental exploration.[3] In the present study, the definitions orally delivered to the subjects contained a detailed description of the structural properties of the items, permitting the construction of complex visual images. On the contrary, in the work of D'Esposito *et al.*, the subjects had to generate a different mental image every second, while listening to a concrete word.[9] The absence of any activation in the right inferior temporal gyrus could then be explained by the fact that the subjects did not have enough time for generating a complex image.

Meanwhile, the left inferior temporal gyrus and the adjacent fusiform gyrus seem to be engaged in the perception of shapes for which there does exist a lexical entry, i.e. that are verbalizable.[20,21] This also seems to be true in the field of visual imagery : the mental visualization of objects serving as landmarks for mental exploration,[3,22] or of objects or animals after listening to their name[9] or the mental imagery of letters of the alphabet,[4] implicate the left inferior temporal gyrus. Moreover, apart from its implication in visual imagery of object, this region seems specifically implicated in language processing and is part of a larger area that has been referred to as the basal temporal language area.[23] As a matter of fact, low intensity electrical stimulations in this cortical territory have elicited naming deficits with preserved identification capacity of either visually presented or auditorily described objects,[23,24] supporting the view that this left infero-temporal area is critically involved in the interaction between language and visual representations.

The lateralization of activations in the inferior temporal cortex could vary, then, according to two distinct dimensions: the presence or the absence of right temporal cortex activations would depend on the complex or simple nature of the visual mental image, while the left activations would depend on it being or not verbalizable. The presence of bilateral activation in previous studies[3,4] as well as in the present work would then be due to the simultaneous presence of these two properties.

Absence of activation in the lateral occipital and the primary visual area (PVA): In the present study, the lateral part of the occipital lobe showed no activation in the CONC condition, whereas activations in the superior or middle occipital gyri were repeatedly reported during spatial imagery tasks.[3,6,22] The absence of lateral activation in the present protocol could be due to an interaction between the two main components of the CONC task, namely sustained listening to and understanding spoken language and simultaneous visual mental imagery activity. As a matter of fact, several studies have reported rCBF decreases in occipital visual areas when the mental image was generated from verbal instructions.[10,11] In a previous study, however, we reported a bilateral rCBF increase in the superior occipital gyrus during the mental construction of a structure described verbally.[3] In the present study, it seems thus likely that the absence of occipital activation was related to the nature of the task which asked mainly for a treatment of the figural properties of the evoked objects. One could then hypothesize that visual mental imagery of objects is a function of the inferior temporal and fusiform gyri, belonging to the ventral pathway, rather than of associative visual areas located at the lateral part of the occipital lobe as is the case with spatial mental imagery.

No activation was detected in PVA in the CONC condition when compared either to the REST or to the ABST condition, a finding to be added to the ongoing debate concerning the role played by this region in visual mental imagery.[7,25,26] Within this framework, the present result demonstrates that, as previously demonstrated for spatial mental imagery,[3,6,22] PVA is not necessarily implied in object visual imagery.

Visual working memory and visual mental imagery: One can assume that the generation, the upholding and the transformation of visual mental images requires the engagement of the visual working memory. In our study, we think that its implication during the CONC condition is reflected by the two activation foci located in the superior frontal gyrus and at the intersection between the middle frontal sulcus and the precentral sulcus. The activation of the superior frontal gyrus, corresponding to the left lateral premotor region known to be engaged in the spatial working memory,[27,28] could express the transformations operated on the mental images when subjects listened to the word definition. In fact, activations in this region have been reported in PET studies where mental images were built up progressively or explored mentally,[3] or in which the subjects were asked to perform a mental navigation.[22] Such activations were, however, absent when no transformation of the image was required[4,5,22]. Similarly, the other frontal activation located in the depth of the middle frontal sulcus, at its intersection with the precentral sulcus could then correspond to the implication of the object working memory,[28] required for upholding and refreshing the mental image of the object or the animal described in the definition.

Consequence of the complexity of the verbal content on the language comprehension network: Activations of the bilateral superior temporal gyri and of the Broca's area were observed in both the CONC and ABST conditions, when compared with REST. This network is known to be implicated in the listening and comprehension of continuous spoken language.[29,30] However, as revealed by the ABST minus CONC subtraction, listening to definitions of abstract words elicited more intensive and more outspread activations of this network. The post-PET recall of these types of words by their definition shows a lower score for abstract words, expressing probably a greater difficulty in understanding the definition of these words as compared to concrete ones. The bilateral activations of the superior temporal gyrus would then be related to the processing of more complex phrases in the course of the ABST condition. In fact, it has been shown that the volume and the intensity of superior temporal activations increased bilaterally with the complexity of spoken phrases.[30] The activation of the right middle temporal gyrus, at the limit of the temporal pole, was somewhat unexpected, although bilateral activation of the temporal poles has been previously described during story listening.[29] The present right middle temporal gyrus activation specific to the ABST condition emphasizes the capabilities of the right hemisphere in spoken language processing.

References

1. Paivio A. *Mental Representations:* A Dual Coding Approach. New York: Oxford University Press, 1986.
2. Denis M and Cocude M. *Mem Cogn* **20**, 497–506 (1992).
3. Mellet E, Tzourio N, Crivello F *et al. J Neurosci* **16**, 6504–6512 (1996).
4. Kosslyn SM, Alpert NM, Thompson WL *et al. J Cogn Neurosci* **5**, 263–287 (1993).
5. Roland PE and Gulyas B. *Cerebr Cortex* **1**, 79–93 (1995).
6. Mellet E, Tzourio N, Denis M and Mazoyer B. *J Cogn Neurosci* **7**, 433–445 (1995).
7. Kosslyn SM, Thompson WL, Kim IJ and Alpert NM. *Nature* **378**, 496–498 (1995).
8. Le Bihan D, Turner R, Zeffiro TA *et al. Proc Natl Acad Sci USA* **90**, 11802–11805 (1993).
9. D'Esposito M, Detre JA, Aguirre GK *et al. Neuropsychologia* **35**, 725–730 (1997).
10. Fink GR, Markowitsch HJ, Reinkemeier M *et al. J Neurosci* **16**, 4275–4282 (1996).
11. Buckner RL, Raichle ME, Miezin FM and Petersen SE. *J Neurosci* **16**, 6219–6235 (1996).
12. Vandenberg SG and Kuse AR. *Percept Motor Skills* **47**, 599–604 (1978).
13. Likert A and Quasha WH. *Revised Minnesota Paper Form Board Test (Series AA).* New York: The Psychological Corporation, 1941.
14. Friston KJ, Holmes A, Worsley K *et al. Hum Brain Map* **2**, 189–201 (1995).
15. Haxby JV, Horwitz B, Ungerleider LG *et al. J Neurosci* **14**, 6336–6353 (1994).
16. Kanwisher N, Woods RP, Iacoboni M and Mazziotta JC. *J Cogn Neurosci* **9**, 133–142 (1997).
17. Martin A, Wiggs CL, Ungerleider LG and Haxby JV. *Nature* **379**, 649–652 (1996).
18. Schacter DL, Reiman E, Uecker A et al. *Nature* **376**, 587–590 (1995).
19. Faillenot I, Sakata H, Costes N *et al. NeuroReport* **8**, 859–862 (1997).
20. Sergent J, Ohta S and MacDonald B. *Brain* **115**, 15–36 (1992).
21. Kosslyn SM, Alpert NM, Thompson WL *et al. Brain* **117**, 1055–1071 (1994).
22. Ghaëm O, Mellet E, Crivello F *et al. NeuroReport* **8**, 739–744 (1997).
23. Lüders H, Lesser RP, Dinner DS *et al. Brain* **114**, 743–754 (1991).
24. Malow BA, Blaxton TA, Sato S *et al. Epilepsia* **37**, 245–252 (1996).
25. Roland PE and Gulyas B. *Trends Neurosci* **17**, 281–286 (1994).
26. Kosslyn SM and Ochsner KN. *Trends Neurosci* **17**, 290–292 (1994).
27. Jonides J, Smith EE, Koeppe RA *et al. Nature* **363**, 623–624 (1993).
28. Courtney SM, Ungerleider LG, Keil K and Haxby JV. *Cerebr Cortex* **6**, 39–49 (1996).
29. Mazoyer BM, Dehaene S, Tzourio N *et al. J Cogn Neurosci* **5**, 467–479 (1993).
30. Just MA, Carpenter PA, Keller TA *et al. Science* **274**, 114–116 (1996).

ACKNOWLEDGEMENTS: The authors are deeply indebted to their colleagues V. Beaudouin, P. Lochon, O. Thirel and M. Marie for their invaluable help in tracer production and data acquisition. E.M. has been supported in part by a grant from the Fondation pour la Recherche Médicale.

Received 12 November 1997; accepted 11 January 1998