



# A neural basis for category and modality specificity of semantic knowledge

S.L. Thompson-Schill<sup>a,\*</sup>, G.K. Aguirre<sup>b</sup>, M. D'Esposito<sup>b</sup>, M.J. Farah<sup>a</sup>

<sup>a</sup> Department of Psychology, University of Pennsylvania, Philadelphia PA, U.S.A.

<sup>b</sup> Department of Neurology, University of Pennsylvania, Philadelphia PA, U.S.A.

Received 9 March 1998; accepted 24 August 1998

## Abstract

Prevalent theories hold that semantic memory is organized by sensorimotor modality (e.g., visual knowledge, motor knowledge). While some neuroimaging studies support this idea, it cannot account for the category specific (e.g., living things) knowledge impairments seen in some brain damaged patients that cut across modalities. In this article we test an alternative model of how damage to interactive, modality-specific neural regions might give rise to these categorical impairments. Functional MRI was used to examine a cortical area with a known modality-specific function during the retrieval of visual and non-visual knowledge about living and non-living things. The specific predictions of our model regarding the signal observed in this area were confirmed, supporting the notion that semantic memory is functionally segregated into anatomically discrete, but highly interactive, modality-specific regions. © 1999 Published by Elsevier Science Ltd. All rights reserved.

*Keywords:* Fusiform gyrus; fMRI; Imagery; Semantic memory; Visual association cortex; Neuroimaging

## 1. Introduction

Several alternatives have been offered to explain the existence of selective deficits in the knowledge of either living [9, 19] or non-living things [17, 18]. The most straightforward interpretation of category-specific deficits is that semantic memory is organized by taxonomic category (e.g., living things). Neuroimaging studies have supported the idea that these impairments indicate the existence of category-specific brain regions [6, 15]. An alternative account holds that semantic knowledge may be organized into different sensorimotor modalities which reflect the origin or form of the information (e.g., visual knowledge) [2, 19]. The contribution of knowledge from different modalities to the representation of an object will vary, depending on the defining or distinguishing features of the object. The representations of living and non-living things differ in terms of the relative proportion of features from visual and non-visual modalities, with living things represented by a larger proportion of visual features than non-living things (Exp. 1) [8]. Thus, the loss of information about visual knowledge would disproportionately impact knowledge about living things [19].

The interpretation of category-specific knowledge deficits within the framework of modality-specific semantic representations has the advantage of parsimony, in that it reflects the sensorimotor organization that is already known to exist in the brain. Additionally, there is also fMRI evidence that supports a modality-specific organization of semantic knowledge [11]. However, a simple modality-specific hypothesis fails to account for one aspect of category-specific knowledge impairments: deficits have been reported for both visual and non-visual knowledge of living things [4, 7, 13, 14, 19]. Whereas lesions of a modality-specific area would be expected to disrupt only knowledge related to that modality, in fact, categorical impairments can occur across all modalities.

How might cross-modal, category-specific impairments occur following lesions to modality-specific areas? A modality-specific hypothesis can account for these impairments when it is taken together with the idea of active, distributed mental representations that depend on collateral support throughout the representation [3]. If knowledge representations are distributed and interactive, with each part of an object's representation providing collateral activation to the other parts, damage to one part may affect the ability of remaining intact parts to become activated [8]. For categories of knowledge with a predominant modality of representation, retrieval of information about any modality may require retrieval

\* Corresponding author. Tel.: +1 215 614-1976; fax: +1 215 349-8464; e-mail: sschill@psych.upenn.edu

from the primary modality. For example, lesions to a visual-modality specific area could produce deficits across modalities for knowledge of living things by such a mechanism. In other words, retrieval of visual information would be obligatory whenever knowledge about living things in general is queried. This interactive modality-specific organization of semantic knowledge has been described in a computational model [8] and is consistent with neuropsychological data [7, 13, 14, 19].

If the interactive modality-specificity hypothesis is correct, then this differential recruitment should be reflected in areas which subserve retrieval of visual information. Activity in the left fusiform gyrus has been observed during both the generation of color words [11] and mental imagery [5], suggesting the involvement of this region in the retrieval of visual information, even in the absence of overt visual stimulation. We examined the dependence of left fusiform activity, and by implication visual knowledge, upon retrieval of information regarding living and non-living things.

We studied five neurologically normal, young subjects (4 males, 1 female) with functional magnetic resonance imaging (fMRI). We asked each subject yes/no questions about either visual or non-visual characteristics of living and non-living things. This design allowed us to manipulate the modality of semantic retrieval (visual or non-visual) independently of the category of object being retrieved (living or non-living). In addition, we included a non-semantic baseline task in which subjects listened to nonsense auditory stimuli. There are two critical predictions of the interactive modality-specificity hypothesis which can be tested within this setting. First, activity should be present in the left fusiform gyrus during retrieval of non-visual information about living things. This is a straightforward prediction of the claim that visual knowledge is obligatory during the retrieval of any information about living things. Second, activity in the left fusiform gyrus should be greater during retrieval of visual information relative to non-visual information for non-living things. This prediction is based on the assumed (and empirically-demonstrated [8]) differences in the proportion of visual knowledge in the representations of living and non-living things. Failure to observe either of these effects would tend to dispute the interactive modality-specificity hypothesis.

## 2. Methods

### 2.1. Cognitive task

We asked yes/no semantic memory questions about the 48 living things (e.g., animals, flowers, vegetables, etc.) and 48 non-living things (e.g., clothing, furniture, kitchen utensils, etc.) used by Warrington and Shallice [19]. For each item, we asked a question about either

visual or non-visual characteristics, yielding four types of questions as in the following examples:

(i) *Visual–Living*

Does a parrot have a curved beak?

Do ducks have long ears?

(ii) *Non-visual–Living*

Are snails edible?

Are pandas found in China?

(iii) *Visual–Non-living*

Are both front and back ends of a submarine approximately the same width?

Are bows of violins longer than violins?

(iv) *Non-visual–Non-living*

Does a toaster use more electricity than a radio?

Can headphones play stereo music?

The questions were selected from a larger set of stimuli described by Farah et al. [7], and they were matched in terms of accuracy and, for visual and non-visual questions, in terms of response latency. (The duration of the questions about living things was, however, shorter than for non-living things, resulting in a shorter response time from the onset of the trial. While having no effect on the analyses we report here, this discrepancy precluded direct comparisons of living and non-living things.) Only one question, either visual or non-visual, was asked about a given object; across scans the type of question asked about each object was counterbalanced.

Questions were digitally recorded and presented by a Macintosh computer using PsyScope software to headphones on the subject. The stimulus onset asynchrony was 5 s and the duration of each question ranged from 2–4 s. Subjects indicated a yes-no response on a keypad. For baseline trials, the auditory stimuli were digitally-reversed to create stimuli which contained no semantic information but which were matched to test stimuli in terms of acoustic properties. This was performed for living and non-living trials separately to match duration, yielding two baseline periods. In the baseline task, subjects also made a keypad response during the trial.

Trials were organized into 30 s blocks of six questions from one of the six trial types (four types of questions and two baselines). The presentation order of the blocks was fixed so that a subject cycled through each of the six trial types twice during a 6 m scan. There were two scans, yielding four blocks of each of the question types and eight baseline blocks. Across subjects the presentation order of the blocks was counterbalanced.

### 2.2. Image acquisition

Following the acquisition of sagittal and axial T1 weighted localizer images, gradient echo, echoplanar fMRI was performed in 16 contiguous 5 mm axial slices

(TR = 2000 ms, TE = 50 ms,  $64 \times 64$  pixels in a 24 cm field of view) using a 1.5 T GE signa system equipped with a fast gradient system and the standard quadrature head coil. Twenty seconds of ‘dummy’ gradient and RF pulses preceded the actual data acquisition to approach tissue steady state magnetization. Head motion was minimized using foam padding. Two, 6 m scans were conducted for each subject, resulting in 360 observations per voxel per subject.

### 2.3. Image analysis

Off-line data processing was performed on SUN Sparc workstations using programs written in Interactive Data Language (Research Systems, Boulder, CO, U.S.A.). After image reconstruction, a slice-wise motion compensation method was utilized which removed spatially coherent signal changes via the application of a partial correlation method to each slice in time. The raw data for each subject were transformed to a standardized spatial frame [16] by landmark guided, nine-parameter differential scaling and spatially smoothed by convolution with a 5 voxel FWHM Gaussian kernel. Spatial smoothing was undertaken to account for residual differences in anatomy following realignment.

Voxel-wise analysis was performed using a general linear model for serially correlated error terms [21]; included within the model was an estimate of intrinsic temporal autocorrelation [1], global signal covariates, and sine and cosine regressors for frequencies below that of the task. The global signals were not significantly correlated with the selection comparison for any of the tasks. Temporal data were smoothed with an empirically derived estimate of the hemodynamic response of the fMRI system [22]. This analysis has been empirically demonstrated to hold the map-wise false positive rate at or below tabular values [1]. A critical  $t$ -value of 3.8 and a cluster requirement of 14 voxels ( $1 \text{ cm}^3$ ) was used for each map to control  $\alpha$  at 0.05 [10, 20]. To test the interaction between category and modality, a  $3 \text{ cm}^3$  region of interest (ROI) was centered upon the local maxima of activation in the fusiform gyrus ( $-33, -48, -18$ ) from the prior mental imagery experiment [5]. The interaction was tested for the average time series in this ROI, and was evaluated with an uncorrected  $\alpha = 0.05$ .

## 3. Results and discussion

We examined activity in ventral occipito-temporal cortex during each of the four retrieval conditions. Within this area (in addition to other cortical regions, see Table 1), we found increased activity relative to baseline in all four conditions. In all four conditions, increased activity was observed in the left middle temporal gyrus (BA21, Fig. 1a). For three of the four conditions—visual ques-

tions about living and non-living things and non-visual questions about living things—this area of increased activity extended ventrally into the fusiform gyrus (BA 37). In a region of interest within the fusiform gyrus (defined around  $-33, -48, -18$  on the basis of previous studies [5]), the pattern of activity we observed depended upon both the modality and the category being retrieved, as evidenced by a significant interaction,  $t = 1.9$ ,  $P < 0.05$ . (Fig. 1b). The responses of this ventral region were of particular interest given our assumption that the fusiform gyrus plays a role in visual semantic retrieval.

To characterize the interaction observed in the fusiform gyrus, we directly compared visual and non-visual questions separately for living and non-living things (Fig. 2). As predicted, for non-living things increased activity in the left fusiform gyrus (local maximum:  $-38, -56, -11$ ) was associated with visual retrieval. For living things, no differences in the left fusiform gyrus, or elsewhere, were associated with visual retrieval. Even with a more liberal mapwise  $\alpha = 0.15$ , no differences between visual and non-visual knowledge retrieval were observed for living things. These findings confirm that the retrieval of visual knowledge about non-living things is greater during task conditions in which visual knowledge is specifically required, but that comparable visual knowledge is retrieved about living things regardless of the task conditions.

As predicted by the interactive modality-specificity hypothesis, activity in the fusiform gyrus, assumed to reflect visual semantic retrieval, was found to be dependent upon modality of retrieval and category of retrieval. Increased fusiform activity was (1) present during retrieval of non-visual knowledge of living things compared to baseline; and, (2) was modulated by modality of retrieval only for non-living things. While these findings do not rule out the possibility of category-specific neural regions, they do provide a neural basis for the interactive modality-specificity hypothesis. That is, we have identified a region of the brain that exhibits changes in activity that reflect an interactive yet modality-specific role in semantic knowledge. Heretofore, consideration of modality-specific bases of cross-modal, category-specific knowledge deficits was limited to theoretical arguments [7] and sufficiency proofs using computational demonstrations [8].

An assumption of the current study that was necessary to test this hypothesis was that activity in the left fusiform gyrus could be assumed to measure visual semantic retrieval. This assumption was based on the results of previous neuroimaging studies that measured brain activity during retrieval of visual information [11] and mental imagery [5]. Consistent with this assumption, we found differences between visual and non-visual knowledge retrieval in the left fusiform gyrus, though of course this does not constitute an independent check upon this assumption. As a more rigorous test of this assumption,

Table 1

Local maxima of statistical maps of the comparison of each semantic task vs the non-semantic baseline task. Numbers in parentheses refer to Brodmann's areas (BA). Local maxima which fall within the fusiform gyrus (BA 37) are indicated in boldface. Co-ordinates are expressed in millimeters in the Talairach and Tournoux brain atlas:  $x$ , medial-lateral axis (negative, left);  $y$ , anterior-posterior axis (negative, posterior);  $z$ , dorsal-ventral axis (negative, ventral). L, left; R, right

Brain region	Category																
	Living				Non-living												
	Visual		Non-visual		Visual		Non-visual										
	Tal. ( $x,y,z$ )	(BA) $t$	Tal. ( $x,y,z$ )	(BA) $t$	Tal. ( $x,y,z$ )	(BA) $t$	Tal. ( $x,y,z$ )	(BA) $t$									
Temporal	<b>-41</b>	<b>-53</b>	<b>-11 (L37)</b>	<b>8.81</b>	<b>-45</b>	<b>-45</b>	<b>-8 (L37)</b>	<b>4.69</b>	<b>-41</b>	<b>-53</b>	<b>-11 (L37)</b>	<b>10.69</b>					
													-49	-45	-4 (L21)	8.80	
Frontal	41	-34	4 (R22)	4.66					49	-30	4 (R22)	7.77		53	-30	4 (R22)	8.22
	-34	19	-4 (L47)	6.36	-34	15	-4 (L47)	4.24	-30	19	-4 (L47)	8.14		-30	19	0 (L47)	5.73
	-41	-4	34 (L6)	7.26					-38	-4	34 (L6)	10.69		-38	-4	38 (L6)	7.13
Cingulate	-11	0	41 (L24)	4.61										11	4	34 (R24)	5.81

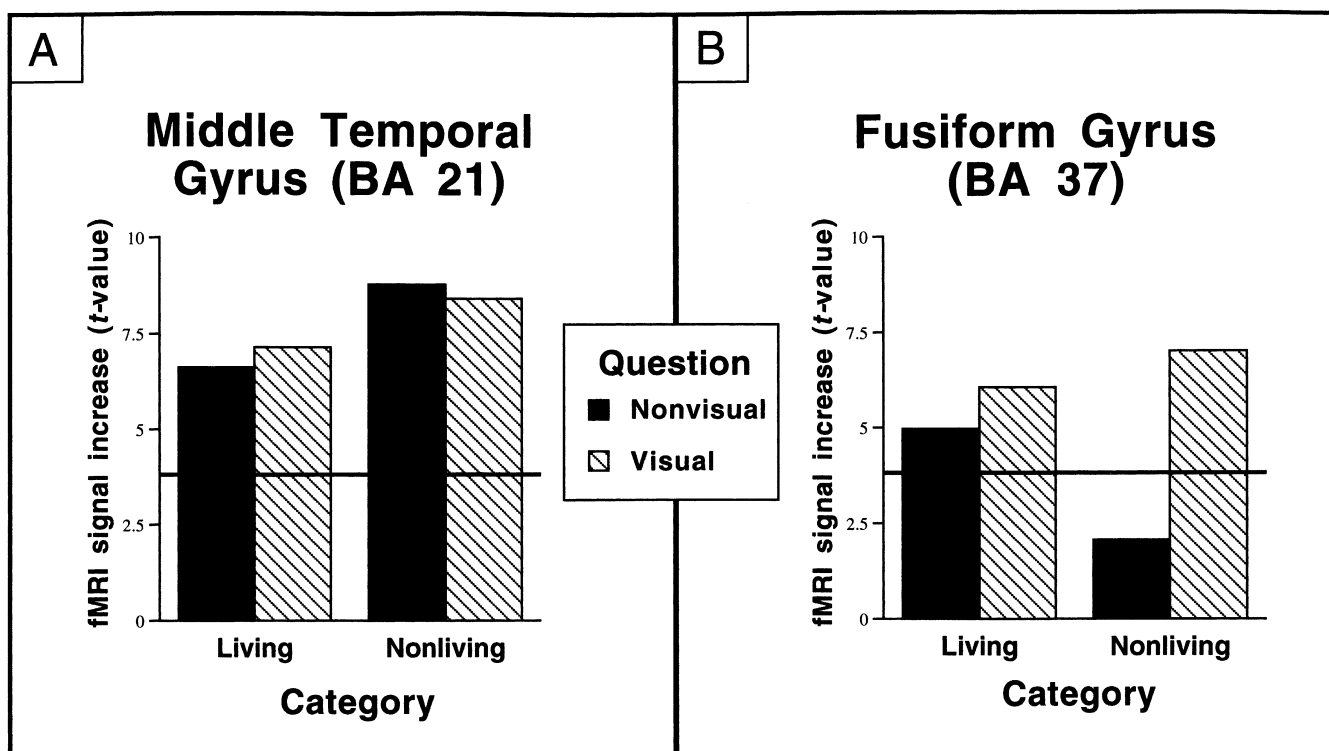


Fig. 1. (A) At a region in the left, middle temporal gyrus ( $-49$ ,  $-45$ ,  $-4$ ),  $t$ -values from the comparison of each condition against baseline, showing significant increases in fMRI signal for all four conditions. (B) At a region in the left fusiform gyrus ( $-33$ ,  $-48$ ,  $-18$ ) identified in a prior study on mental imagery [5],  $t$ -values from the comparison of each condition against baseline, showing significant increases in fMRI signal for visual and non-visual questions about living things and for visual questions (but not non-visual questions) about non-living things. In both figures, the horizontal bar indicates the critical value ( $t = 3.8$ ) for comparisons vs baseline with  $\alpha = 0.05$ .

we examined a single subject from the present study who had also participated in the mental imagery experiment [5]. In that experiment, we found increased activation in the left fusiform gyrus (local maximum:  $-33$ ,  $-48$ ,  $-18$ )

during a task requiring active image generation relative to passive audition of abstract words. We were able to directly examine the co-localization of cortical regions associated with mental imagery and visual semantic

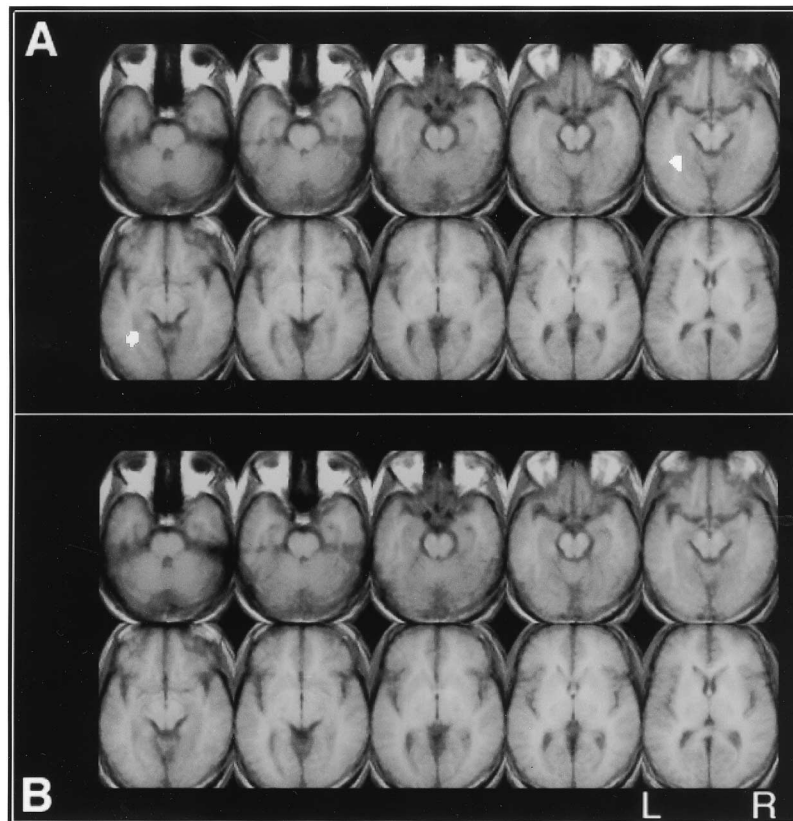


Fig. 2. 3.75 mm axial slices through ventral temporo-occipital cortex ( $z = -26$  to  $z = 8$ ) showing voxels that exceeded a threshold of  $t = 3.8$ , mapwise  $\alpha = 0.05$  (corrected for multiple comparisons) in the comparison of retrieval of visual vs non-visual knowledge ( $N = 5$ ). (A) For non-living things, greater activity in the fusiform gyrus (local maximum:  $-38, -56, -11$ ) was observed during the retrieval of visual knowledge than non-visual knowledge. (B) For living things, no differences between visual and non-visual knowledge retrieval were found in the fusiform gyrus or elsewhere.

retrieval in this subject. During both tasks, this subject showed increased activity at the same point in the left fusiform gyrus (Fig. 3). The consistency of the regions identified both at the level of the group analyses and also in an individual subject suggest that both retrieval of visual knowledge and active image generation involve the same cognitive processes.

Despite these findings, one might choose to question our initial assumption that the fusiform gyrus subserves visual retrieval. One possible criticism is that the activity in this area seen in the present and previous studies is the result of some cognitive process which is consistently confounded with visual retrieval, and not visual retrieval per se. Even if this were the case, however, activation in the fusiform gyrus would nonetheless provide a reliable index of visual retrieval, allowing the inferences made in the present report.

Although previous studies of visual knowledge retrieval have implicated the left fusiform gyrus, neuroimaging studies of category specificity have not. For example, Martin and colleagues [12] reported activation in the medial occipital cortex for naming animals rela-

tive to naming tools. It is conceivable that the process of retrieving an object's name from a picture of that object would require visual semantic knowledge for both living and non-living things, as we observed when asking visual questions of both types of things. Under these conditions, one would not expect to find evidence of category-specificity in the fusiform gyrus during object naming.

The reliance on assumptions about functional localization to make inferences about psychological processes represents a fundamental difference from the types of inference common to neuroimaging studies. Traditionally, neuroimaging has been used to infer the neural mechanisms of a well-characterized process. In the work presented here, the direction of this inference was inverted. Using our knowledge regarding the function of a particular neuroanatomical area, the fusiform gyrus, we were able to draw inferences about an ambiguous cognitive process. Future demonstrations of unambiguous structure–function relationships will increasingly allow us to treat cognitive as the dependent variable and use neuroimaging to inform cognitive theory.

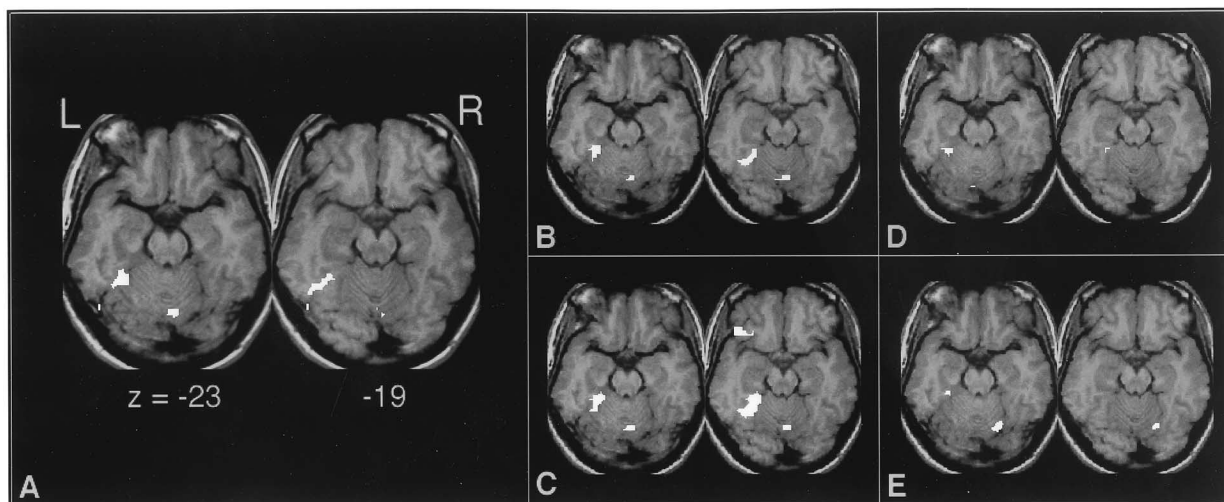


Fig. 3. 3.75 mm axial slices (in standard space) through the fusiform gyrus ( $z = -23$  to  $-19$ ) of a single subject showing voxels that exceeded a threshold of  $t = 3.5$ . (A) During an explicit mental imagery task (previously reported [5]), greater activation was found in the left fusiform gyrus (local maximum:  $-38, -49, -23$ ) during mental imagery of concrete words than during passive audition of abstract words. In the present study, performed two years later, the same subject showed increased activity in the left fusiform gyrus during retrieval of (B) visual knowledge about living things ( $-26, -42, -24$ ), (C) visual knowledge about non-living things ( $-26, -38, -24$ ), (D) non-visual knowledge about living things ( $-26, -45, -25$ ), (E) non-visual knowledge about non-living things ( $-33, -59, -14$ ). As seen in (D), fusiform activation was present in the retrieval of visual knowledge about living things, as predicted by the interactive modality-specificity hypothesis.

## Acknowledgements

We thank Glyn Humphreys and an anonymous reviewer for helpful comments on an earlier version of this manuscript. This research was supported by the James S. McDonnell Foundation and by National Institutes of Health Grants AG05743, NS34030, AG14082, AG00756, NS01762, and AG13483.

## References

- [1] Aguirre GK, Zarahn E, D'Esposito M. Empirical analyses of BOLD fMRI statistics II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *NeuroImage* 1997;5:199–212.
- [2] Allport DA. Distributed memory, modular subsystems and dysphasia. In: Newman SK, Epstein R, editors. *Current Perspectives in Dysphasia*. Edinburgh: Churchill Livingstone, 1985. p. 207–44.
- [3] Anderson JA, Silverstein JW, Ritz SA, Jones RS. Distinctive features, categorical perception and probability learning: Some applications of a neural model. 1997;84:000–000.
- [4] Caramazza A, Shelton JR. Domain-specific knowledge systems in the brain: The animate-inanimate distinction. *Journal of Cognitive Neuroscience* in press.
- [5] D'Esposito M, Detre JA, Aguirre GK, Alsop DC, Tippet LJ, Farah MJ. A functional MRI study of mental image generation. *Neuropsychologia* 1997;35:725–30.
- [6] Damasio H, Grabowski TJ, Tranel D, Hichwa RD, Damasio AR. A neural basis for lexical retrieval. *Nature* 1996;380:486–96.
- [7] Farah MJ, Hammond KH, Mehta Z, Ratcliff G. Category-specificity and modality-specificity in semantic memory. *Neuropsychologia* 1989;27:193–200.
- [8] Farah MJ, McClelland JL. A computational model of semantic memory impairment: Modality-specificity and emergent category-specificity. *Journal of Experimental Psychology: General* 1991;120:339–57.
- [9] Farah MJ, McMullen PA, Meyer MM. Can recognition of living things be selectively impaired? *Neuropsychologia* 1991;29:185–93.
- [10] Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta J, Evans AC. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping* 1994;1:210–20.
- [11] Martin A, Haxby JV, Lalonde FM, Wiggs CL, Ungerleider LG. Discrete cortical regions associated with knowledge of color and knowledge of action. *Science* 1995;270:102–5.
- [12] Martin A, Wiggs CL, Ungerleider LG, Haxby JV. Neural correlates of category-specific knowledge. *Nature* 1996;379:649–52.
- [13] Mehta Z, Newcombe F, De Haan E. Selective loss of imagery in a case of visual agnosia. *Neuropsychologia* 1992;30:645–55.
- [14] Silveri MC, Gainotti G. Interaction between vision and language in category-specific semantic impairment. *Cognitive Neuropsychology* 1988;5:677–709.
- [15] Spitzer M, Kwong KK, Kennedy W, Rosen BR, Belliveau JW. Category-specific brain activation in fMRI during picture naming. *NeuroReport* 1995;6:2109–12.
- [16] Talairach J, Tournoux P. *Co-planar stereotaxic atlas of the human brain*. Thieme Medical Publishers, Inc., 1988.
- [17] Warrington EK, McCarthy R. Category specific access dysphasia. *Brain* 1983;106:859–78.
- [18] Warrington EK, McCarthy R. Categories of knowledge: Further fractionations and an attempted integration. *Brain* 1987;110:1273–96.
- [19] Warrington EK, Shallice T. Category specific semantic impairments. *Brain* 1984;107:829–54.
- [20] Worsley KJ. Local maxima and the expected euler characteristic of excursion sets of chi-squared,  $f$  and  $t$  fields. *Advanced Applied Probability* 1994;26:13–42.
- [21] Worsley KJ, Friston KJ. Analysis of fMRI time-series revisited—again. *NeuroImage* 1995;2:173–82.
- [22] Zarahn E, Aguirre GK, D'Esposito M. Empirical analyses of BOLD fMRI statistics I. Spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage* 1997;5:179–97.