

www.elsevier.com/locate/ynimg NeuroImage 30 (2006) 992 - 1002

Semantic relevance explains category effects in medial fusiform gyri

Andrea Mechelli,^{a,*} Giuseppe Sartori,^b Paola Orlandi,^b and Cathy J. Price^c

^aInstitute of Psychiatry, King's College London, 103 Denmark Hill, London SE5 8AF, UK

^bDepartment of Psychology, Università di Padova, Via Venezia 8, 35100 Padova, Italy

^cWellcome Department of Imaging Neuroscience, Institute of Neurology, 12 Queen Square, London WC1N 3BG, UK

Received 1 September 2005; revised 10 October 2005; accepted 14 October 2005 Available online 15 December 2005

We used functional Magnetic Resonance Imaging to explore the neural correlates of semantic relevance in 12 healthy participants performing a picture-naming task. In addition, we tested the hypothesis that category effects typically found in functional imaging can be partly explained in terms of different semantic relevance for animals and artefacts. We report that semantic relevance modulates neuronal responses in the medial fusiform gyrus bilaterally. As predicted, category effects in this region are strongly modulated by the semantic relevance of the items. Specifically, the effect of artefacts > animals is greatly reduced when the two categories are matched for semantic relevance. Thus, the present study demonstrates that neuronal responses during concept retrieval are modulated by the semantic relevance of the features. It also suggests that increased activation in the medial fusiform gyrus typically found for artefacts > animals can be explained by different semantic relevance for animal and artefact items.

© 2005 Elsevier Inc. All rights reserved.

Introduction

The question of how concepts are represented in the human brain is still highly debated. The most popular theory suggests that conceptual representations are based on semantic features (e.g., sensory and non-sensory). Recently, it has been proposed that conceptual representations depend on the relative contribution of their semantic features to the meaning of a concept (e.g., semantic relevance). The present study used functional Magnetic Resonance Imaging (fMRI) to investigate the neural correlates of semantic relevance during concept retrieval, and furthermore to explore how these effects interact with category.

One way of analysing semantic features involves grouping them according to their content. In this regard, one of the most frequently investigated distinctions is that between Sensory and

* Corresponding author. Fax: +44 20 78131420.

E-mail address: a.mechelli@iop.kcl.ac.uk (A. Mechelli). Available online on ScienceDirect (www.sciencedirect.com). Non-sensory features. Consider for example the concept Dog.^{1,2} Sensory features include "has four legs", non-sensory features include functional (e.g., "is used for hunting"), associative (e.g., "likes to chase cats") and encyclopaedic features (e.g., "may be one of many breeds").^{3,4} This proposal has been enormously influential, spanning an entire area of empirical enquiry (Warrington and Shallice, 1984; Allport, 1985; Warrington and McCarthy, 1987; Farah and McClelland, 1991; Saffran, 2000; Martin and Chao, 2001; Thompson-Schill et al., 1999; Rogers et al., 2005). A number of functional neuroimaging studies have investigated the neural basis of semantic features by directly comparing sensory and non-sensory tasks (Martin et al., 1995; Gauthier et al., 2000; Thompson-Schill, 2003; Noppeney and Price, 2003; Hauk et al., 2004; Kellenbach et al., 2005). Other functional imaging studies have reported differential activation for different categories of object stimuli (Martin et al., 1996; Mummery et al., 1998; Chao et al., 1999; Tyler et al., 2000; Devlin et al., 2002). Although there is a growing consensus that both action tasks and tool stimuli increase activation in a distributed left hemisphere visuo-motor system, many inconsistent results have also been reported. These inconsistencies could be due to the fact that concepts in our lexicon vary in relation to many other factors (Cree and McRae, 2003), including frequency, familiarity, age-of-acquisition etc. Most critically these dimensions differ for Living and Non-living (Cree and McRae, 2003).

^{1053-8119/\$ -} see front matter @ 2005 Elsevier Inc. All rights reserved. doi:10.1016/j.neuroimage.2005.10.017

¹ Concept names are printed in italics, and names of semantic features in angled brackets.

² Semantic features are also sometimes termed "properties" or "attributes".

³ Throughout this paper, the term "concept" refers to a set of weighted semantic features; semantic feature is used to describe any type of statement about the concept (both Sensory and Non-sensory).

⁴ Functional features are defined in different ways. Some authors use this term for features that directly refer to functions (e.g. "gives milk") others denote physically defined features defined by motor properties (e.g. "used to cut", Farah and McClelland, 1991). Others have defined functional knowledge by exclusion to denote any property that is not physically defined (Thompson-Schill et al., 1999).

Another factor that may be uncontrolled across object categories is semantic relevance. A concept may have uncountable semantic features although those really useful in distinguishing it from closely related concepts may not be numerous. Among dimensions proposed as descriptors of semantic features we can list dominance (Ashcraft, 1978), distinctiveness (Garrard et al., 2001) and most recently semantic relevance (Sartori and Lombardi, 2004; Sartori et al., 2005). Semantic relevance is a measure of the contribution of semantic features to the "core" meaning of a concept. For example, "has a trunk" is a semantic feature of high relevance for the concept *Elephant*, because most subjects use it to define *Elephant*, whereas very few use the same feature to define other concepts. Instead "Has 4 legs" is a semantic feature with low relevance for the same concept, because few subjects use it to define *Elephant* but do use it to define many other concepts. When a set of semantic features is presented, the concept with the highest summed relevance is the one which will be retrieved.

Semantic relevance is the result of two components (see Methods for details). A local component, which measures the importance of the semantic features for the concept, may be interpreted as dominance.⁵ A global component, which measures the importance of the same semantic feature for all the other concepts in the lexicon, may be interpreted as distinctiveness.⁶ While both dominance and distinctiveness alone do not predict accuracy in a "naming-to-description" task, they are highly correlated with naming accuracy when combined into relevance (Sartori et al., 2005). Sartori et al. (2005) also showed that: (i) relevance is better at predicting naming accuracy in a "naming-to-description" task than a number of other parameters such as Age-of-Acquisition, frequency, familiarity and typicality; (ii) relevance is a robust measure, not significantly influenced by the number of concepts in the database or by sampling errors. Critically, living and non-living items differ in terms of their semantic relevance (Sartori and Lombardi, 2004). Specifically, living items may have either low, medium or high semantic relevance. In contrast, non-living items can only have low or medium relevance. This led to the prediction that semantic relevance may account for category-specific deficits in neuropsychological patients. Sartori and Lombardi (2004) found that, when concepts belonging to Living and Non-living categories are equated for relevance, a previously well-established categoryspecific deficit disappears. Furthermore, the selective impairments could be completely reversed, in a naming-to-description task, by presenting stimuli with properly selected levels of relevance (Sartori and Lombardi, 2004). In short, semantic relevance is the best predictor of naming accuracy in healthy subjects and furthermore appears to account for category-specific deficits in neuropsychological patients.

Based on the findings of the behavioural studies in healthy subjects and neuropsychological patients, one would expect that semantic relevance affects brain responses. However, the neural correlates of semantic relevance are currently unknown. The first aim of the present study was therefore to investigate the effect of semantic relevance on brain activation during concept retrieval. For this purpose, we measured brain responses in 12 healthy volunteers who overtly named pictures of items with low, medium or high semantic relevance. Because semantic relevance is based on high order visual features,⁷ we expected its effect to be most apparent within the human ventral visual pathway.

The observation that non-living items typically have greater semantic relevance than living-items raises the possibility that the category effects reported in functional imaging studies may be partly explained in terms of semantic relevance. This hypothesis is supported by the predictive value of relevance in neuropsychological studies of patients with category-specific disorders (Sartori and Lombardi, 2004). The second aim of the present study was therefore to test the hypothesis that category effects within the human ventral visual pathway are modulated by the semantic relevance of the items. In order to test this hypothesis, we manipulated the category of the items which could be either animals or artefacts. We predicted that category effects within the ventral visual pathway would be modulated by the semantic relevance of the items. Specifically, we expected that differences between artefacts and animals to be greatly reduced when the two categories are matched for semantic relevance.

Materials and methods

Relevance of semantic features

Relevance is a measure of the contribution of semantic features to the "core" meaning of a concept. A few semantic features of high relevance are sufficient for an accurate retrieval of the target concept. When a set of semantic features is presented, their overall relevance results from the sum of the individual relevance values associated with each of the semantic features. Relevance may be interpreted as a non-linear combination of dominance (also called production frequency; Cree and McRae, 2003) and distinctiveness (Sartori et al., 2005).

Dominance

Dominance is a measure of how frequently a semantic feature is used in defining a concept (Ashcraft, 1978; Garrard et al., 2001). It is defined as the number of times subjects report a given feature in defining a concept;

Dominance_{*ij*} = number of *times* feature *i* is reported for concept *j*

Distinctiveness

The distinctiveness of a given feature is defined as the inverse of sharedness, which, in turn, is defined as the number of concepts in which the semantic feature appears, divided by

⁵ Dominance is a measure of how frequently a semantic feature is used in defining a concept (Ashcraft, 1978).

⁶ Relevance of semantic features is different from distinctiveness. Distinctiveness is a dimension which is not concept-dependent, since scores are high when the feature is found in only a few concepts (Garrard et al., 2001). Instead, the relevance of a given semantic feature varies across different concepts and, in a way, may be considered concept-dependent. For example, the feature "has a beak" has higher relevance for the concept *Duck* than for the concept *Swan* but the same distinctiveness.

⁷ High order visual features refer to features as "is an animal", "has 4 legs" which are derived from verbal descriptions of subjects defining the concepts. In contrast, low order visual features refer to parameters such as size, line orientation, luminance, texture, color and contrast, which have no semantic interpretation (Cavanagh et al., 1990).

the total number of concepts in the lexicon (or database) (Tyler et al., 2000).

Distinctiveness_{*i*} = N/n_i

where N is the total number of concepts analysed and n_j the number of concepts the feature j appears in.

Relevance

Relevance results from a non-linear combination of dominance and distinctiveness

$Relevance_{ii} = dominance_{ii} * log_2(distinctiveness_i)$

Relevance is different from distinctiveness. Consider this hypothetical example based on 100 subjects defining 100 concepts one of which is Tiger. Consider now the following two features: (i) "has black stripes on yellow background" and (ii) "seen yesterday at the zoo". Assume that both are reported in only 1 of the 100 concepts, exactly in Tiger. Accordingly, the two features have the same distinctiveness which is (N/n = 100/1) = 100. Suppose that all 100 subjects list the feature "has black stripes on vellow background" (dominance = 100) but only one subject reports "seen yesterday at the zoo" (dominance = 1). Then the relevance of "has black stripes on yellow background" will be much higher than "seen yesterday at the zoo" for the concept Tiger but these two features will have the same distinctiveness. Intuitively, the importance of "has black stripes on yellow background" in indexing Tiger is much higher then "seen vesterday at the zoo" and this importance is captured by relevance but not by distinctiveness. The logarithmic damping is intended to capture the idea that, for example, the effect of having a feature in 3 concepts rather than in 1 is larger than the effect of having the same feature in 97 rather than in 94 concepts (for details see Sartori et al., 2005).

Estimating semantic relevance of higher order visual features for pictures

When a picture is presented, the visual features that are encoded in the picture are processed. Each of these features has an associated relevance value. The total relevance of the picture results from the sum of the relevance values of the constituent visual features. We estimated the relevance for 62 pictures of animals and 169 pictures of artefacts taken from various stimulus sets (e.g., Snodgrass and Vanderwart, 1980; Dell'Acqua et al., 2000).

Relevance of visual features

Procedure for estimating semantic relevance of visual features:

- (a) Feature listing of visual features. Eighteen subjects were shown a subset of 50/240 pictures. For each picture they were instructed to list the visual features they could identify in it. Eighteen subjects described every picture.
- (b) For each concept, the number of times a given feature was listed was counted. Synonyms and lemma were collapsed into the same semantic feature according to the procedure reported by Cree and McRae (2003).
- (c) A matrix of 231 pictures \times 490 visual features was obtained.
- (d) From the matrix, which reports the number of times each semantic feature was listed for a given picture, the

relevance was computed according to the following formula:

 $Relevance_{ii} = dominance_{ii} * log_2(distinctiveness_i)$

where

Dominance_{*ii*} = number of *times* feature *i* is reported for picture *j*

Distinctiveness_i =
$$N/n_i$$

with N = the total number of pictures and n = the number of pictures the feature *j* appears in.

Relevance of a picture

The relevance of a picture was calculated by summing up all relevance values of visual semantic features that were reported:

Total Visual relevance of picture_i = Σ_i (relevance_{ii})

This is the measure on which the fMRI study on picture naming due to visual features was based.

fMRI study

Subjects

Informed consent was obtained from 12 right handed volunteers (5 males), aged between 21 and 28 (with a mean age of 24), with English as their first language. This group did not include those subjects who were recruited for the estimation of the semantic relevance of the pictures, see above. None of the volunteers reported a history of neurological or psychiatric illness, or disturbances in speech comprehension, speech production, reading or writing.

Experimental paradigm

The experimental paradigm included two tasks: naming and rest. During the naming condition, subjects overtly named a total of 270 black and white drawings of animal and artefact items. Stimuli included 62 animal items and 169 artefact items. The remaining 39 stimuli were fruit, vegetable or body part items that are not relevant to the present study and will not be discussed further. An additional factor was the semantic relevance of the stimuli: artefacts could have low relevance (38 items, range 112.29-182.33; mean 160.4641), medium relevance (108 items, range 185.68-324.38; mean 240.00) or high relevance (23 items, range 327.48-401.13; mean 355.43); animals could have low relevance (31 items, range 105.44-183.43; mean: 154.20) or medium relevance (31 items, range 184.66-325.43; mean: 219.03). A full list of artefact and animal items and their semantic relevance can be found in Appendix. Stimuli of the same category were presented in blocks of 5, with duration of 600 ms and stimulus onset asynchrony of 3055 ms. During the resting condition, subjects viewed a fixation cross in the centre of the screen. A total of 15 resting blocks were used, each lasting 15.275 s. In order to minimise speech-related movements, subjects were instructed to whisper. Vocal responses were recorded using a microphone to identify unsuccessful trials, which were then discarded from the statistical analysis.

Scanning technique

A Siemens 3T scanner was used to acquire a total of 470 T_2^* -weighted echoplanar images with BOLD contrast. Each echoplanar

image comprised 35 axial slices of 2 mm thickness with 1 mm slice interval and 3×3 mm in-plane resolution. Volumes were acquired with an effective repetition time (TR) of 2.275 s/volume and the first six (dummy) volumes of each run were discarded to allow for T₁ equilibration effects. In addition, a T1-weighted anatomical volume image was acquired from all subjects.

Statistical parametric mapping

Statistical parametric mapping was performed using SPM2 software (Wellcome Department of Imaging Neuroscience, London, UK), running under Matlab 6.5 (Mathworks Inc. Sherbon MA, USA). All volumes from each subject were realigned using the first as the reference and resliced with sinc interpolation. The functional images were spatially normalised (Friston et al., 1995a) to a standard MNI-305 template using nonlinear-basis functions. Functional data were spatially smoothed with a 6-mm full width at half maximum isotropic Gaussian kernel, to compensate for residual variability after spatial normalisation and to permit application of Gaussian random field theory for corrected statistical inference.

First, the functional data were analysed in a subject-specific fashion. In order to remove low-frequency drifts, a high-pass filter was used with a cutoff period of 128 s. Each stimulus was modelled by convolving the onset times with a synthetic hemodynamic response function (HRF). The parameter estimates were calculated for all brain voxels using the general linear model, and the contrast images of interest were computed (Friston et al., 1995b). Second, the subject-specific contrast images were entered into an ANOVA to permit inferences at the population level (i.e., a random effects analysis). The *t*-images for each contrast at the second level were subsequently transformed into statistical parametric maps of the Z statistic.

Two separate statistical analyses were performed, which used a parametric and a categorical approach, respectively. In the parametric analysis, semantic relevance was modelled in a parametric fashion for animal and artefact items independently. This analysis allowed us to look at the linear and nonlinear effect of semantic relevance, for each category independently. In the categorical analysis, low, medium and high relevance items were modelled independently within each category. This resulted in 5 experimental conditions of interest: low relevance animals (range 105.44-183.43; mean: 154.20), medium relevance animals (range 184.66-325.43; mean: 219.03), low relevance artefacts (range 112.29-182.33; mean 160.4641), medium relevance artefacts (range 185.68-324.38; mean 240.00), high relevance artefacts (range 327.48-401.13; mean 355.43). This analysis allowed us to compare animal and artefact activations when relevance was unmatched (e.g., all artefacts > all animals) as well as when relevance was equated (e.g., low relevance artefacts > low relevance animals). We report and discuss regions that showed significant effects at P < 0.05 (corrected for multiple comparisons across the whole brain) with an extent threshold for each cluster of 5 voxels.

Results

Accuracy of vocal responses was very high for both animal (97.4%) and artifact (98%) items. Furthermore, there was no significant difference in accuracy between animal and artifact items (P = 0.345). Our presentation of the functional imaging results starts with a brief report of activation for naming all object categories relative to fixation. We then focus on the parametric

effects of semantic relevance. Finally, we test the hypothesis that category-selective activation can be explained in terms of differential semantic relevance for animal and artefact items.

Effect of naming vs. fixation

First, we examined the overall effect of naming at P < 0.05 (corrected for multiple comparisons across the whole brain). Naming animals and artefacts relative to fixation increased neuronal activity in a number of regions including bilateral inferior, middle and superior occipital, middle temporal, inferior frontal and intra-parietal cortex, see Table 1 and Fig. 1. These results are consistent with previous studies of picture naming that identified similar distributed networks (Bookheimer et al., 1995; Murtha et al., 1999; Price et al., 2005).

Parametric effects of semantic relevance

Linear and nonlinear effects of semantic relevance were investigated when data were (i) limited to artefacts only; (ii) limited to animals only; (ii) averaged over artefacts and animals. We detected a positive linear effect of semantic relevance when the stimuli were artefacts, see Table 2 and Fig. 2. This was observed in bilateral medial fusiform, bilateral middle occipital, right lingual gyrus and right cerebellum (P < 0.05 corrected). In these regions, artefacts with higher semantic relevance elicited greater neuronal responses than artefacts with lower semantic relevance. These effects were not observed for animals, even when the statistical

Table 1

Regions that expressed increased activation for naming relative to fixation
NAMING > FIXATION

Anatomical location	x	у	Ζ	Z score
Frontal				
Left precentral	-50	-8	46	8.4
Right precentral	50	-10	38	8.0
Left inferior frontal/insula	-54	$^{-2}$	18	7.4
Right inferior frontal/insula	54	-4	24	7.4
Left middle superior frontal	$^{-2}$	12	46	8.4
Right middle superior frontal	2	12	46	8.1
Temporal				
Left middle temporal	-54	-46	10	6.9
Right middle temporal	66	-34	8	7.8
Left superior anterior temporal	-56	12	-8	7.8
Occipital				
Left middle occipital	-28	-96	2	8.2
Right middle occipital	18	-98	-4	8.0
Left lateral posterior fusiform	-40	-62	-20	8.2
Right lateral posterior Fusiform	46	-62	-18	8.4
Left lateral middle fusiform	-44	-52	-18	7.5
Right lateral middle fusiform	44	-50	-20	7.9
Left superior occipital	-28	-74	26	6.0
Right superior occipital	30	-74	26	7.4
Parietal				
Left intraparietal sulcus	-28	-60	44	6.2
Right intra-parietal sulcus	28	-68	44	6.7
Other				
Left parahippocampal gyrus	-14	-26	-16	7.4
Left superior cerebellum	-40	-48	-28	9.0
Right superior cerebellum	38	-56	-32	8.7

Co-ordinates [x, y, z] are reported in Talairach space. All Z score are significant at P < 0.05 (corrected for multiple comparisons).

threshold was as low as 0.001 uncorrected. Consequently, there was no significant effect of semantic relevance when data from artefacts and animals were averaged together (P > 0.001 uncorrected) but there was an interaction between semantic relevance and category confirming that the effect of semantic relevance was higher for artefacts than animals (left medial fusiform: Talairach co-ordinates x = -28; y = -54; z = -14 Z score: 4.0; right medial fusiform: Talairach co-ordinates x = 30; y = -48; z = -14 Z score: 4.7; see Table 2). Negative or nonlinear effects of relevance were not observed even when lowering the statistical threshold to P < 0.001 (uncorrected).

Categorical effect of living vs. non-living

After identifying the neural correlates of semantic relevance, we tested our hypothesis that category-sensitive activations may be accounted for by differential semantic relevance for animal and artefact items. This was achieved by comparing artefacts and animals when relevance was unmatched (i.e., all artefacts > all animals) as well as when relevance was equated (e.g., low relevance artefacts > low relevance animals).

We first compared the two categories when relevance was unmatched by directly comparing all animals with all artefacts. In other words, artefacts with low, medium or high relevance were contrasted against animals with low or medium relevance. A significant effect of artefacts > animals was detected in the left and right medial fusiform gyri, see Table 3 and Fig. 3. These regions have been reported to respond more to non-living than living items in previous functional imaging studies (Chao et al., 1999). The reverse comparison (animals > artefacts) did not identify any category-effects that survived our statistical threshold of P < 0.05(corrected). However, when lowering the statistical threshold to P <0.001 (uncorrected), we observed a trend in the right lateral fusiform gyrus (Talairach co-ordinates x = 44; y = -50; z = -22; Z score = 3.2) which was consistent with previous studies comparing animals to artefacts (Chao et al., 1999; Rogers et al., 2005), see Fig. 4.

We then compared the two categories when relevance was equated by discarding artefacts with high relevance. In other words, artefacts with low or medium relevance were contrasted against animals with similar relevance. The effect of artefacts > animals was greatly reduced and no longer survived our statistical threshold of P < 0.05 (corrected). However, a trend could still be detected in both left and right medial fusiform when lowering the statistical threshold to P < 0.001 (uncorrected). We then compared animals and artefacts independently for medium and low relevance items. When only medium relevance items were considered, the effect of artefacts > animals was greatly reduced and no longer survived a statistical

Naming > Fixation



Fig. 1. Brain areas that expressed increased activation for naming relative to fixation (thresholded at P < 0.0001 uncorrected).

ab	le	2
ab	le	2

Regions that expressed a positive linear effect of relevance for artefacts, animals and artefacts > animals

Positive linear effe	ect of r	elevanc	e						
Anatomical	x	у	Z	Z score					
location				Artefacts	Animals	Artefacts > Animals			
Left medial fusiform	-28	-52	-14	5.1	n.s.	4.0			
Right medial fusiform	32	-50	-16	5.6	n.s.	4.7			
Left middle occipital	-29	-92	4	5.5	n.s.	n.s.			
Right middle occipital	34	-82	10	5.1	n.s.	n.s.			
Right lingual gyrus	6	-88	-16	5.2	n.s.	n.s.			
Right cerebellum	18	-84	-24	5.1	n.s.	n.s.			

Co-ordinates [x; y; z] are reported in Talairach space. Z scores significant at P < 0.05 (corrected for multiple comparisons) are reported in bold. The remaining Z scores are significant at P < 0.001 (uncorrected). n.s. = not significant at P < 0.001 (uncorrected).

threshold of P < 0.05 (corrected), but a trend could be detected in the medial fusiform bilaterally with a statistical threshold of P < 0.001 (uncorrected). When only low relevance items were considered, the effect of artefacts > animals was no longer significant even when lowering the statistical threshold to P <0.001 (uncorrected).

To summarise, the parametric analysis revealed that semantic relevance affects neuronal responses in the bilateral medial fusiform for artefact but not for animal items. The categorical analysis revealed that the same region expressed increased activation for artefacts relative to animals when relevance is not matched; however, this effect is strongly reduced when relevance is similar for artefacts and animals.

Discussion

The first objective of the present study was to investigate the neural correlates of semantic relevance during concept retrieval. To this purpose, we measured brain responses in 12 healthy volunteers who overtly named pictures of items which had different semantic relevance. We expected the effect of semantic relevance to be most apparent within the human ventral visual pathway, on the basis that relevance is estimated from high order visual features. We report that semantic relevance increases neuronal responses in the left and right medial fusiform gyrus for artefact but not animal items. This effect appears to be specific to semantic relevance. For instance, parallel analyses revealed that the effect of semantic relevance is not affected when visual complexity, Age-of-Acquisition, frequency, familiarity and typicality are modelled as additional variables of interest. We were specifically concerned with visual complexity because this factor might co-vary with visual relevance in some instances. However, when visual complexity was explicitly modelled in the analysis, a positive effect of semantic relevance was still significant in both left (x = -26; y = -48; z = -14; Z score: 5.6) and right (x = 28; y = -44; z = -12; Z score: 5.8) fusiform gyri. In contrast, no effect of visual complexity was found in these regions even when lowering the statistical threshold to P <



Fig. 2. Effect of semantic relevance on brain responses (x = 32; y = -48; z = -14).

0.001 (uncorrected). A direct comparison revealed that the effect of semantic relevance was significantly greater than the effect of visual complexity in both left (x = -28; v = -48; z = -14; Z score = 4.7) and right (x = 34; v = -48; z = -10; Z score = 4.6) fusiform gyri. Thus, the effect of semantic relevance could not be explained by differences in visual complexity. Importantly, the medial fusiform areas that are modulated by semantic relevance correspond to the areas that are more responsive to artefacts than animals (e.g., see Chao et al., 1999). The second objective of the study was to test the hypothesis that category effects within the human ventral visual pathway are modulated by semantic relevance. This was achieved by comparing artefacts and animals when relevance was unmatched (i.e., all artefacts > all animals) as well as when relevance was equated (e.g., low relevance artefacts > low relevance animals). When artefacts and animal items are compared irrespective of semantic relevance, we find a strong effect of artefacts > animals in the very same fusiform regions modulated by semantic relevance However, when semantic relevance is carefully equated between the two categories, the effect of artefacts > animals is greatly reduced. Two questions are critical for the interpretation of these results. First, why was there a positive effect of relevance in the medial fusiform gyrus for artefact but not for animal items? Second, why was the effect of artefacts > animals reduced when relevance was matched for the two categories? We will now address these questions in turn.

With respect to why a positive effect of relevance was detected for artefact but not for animal items, we considered the possibility that this might be explained by differences in statistical power due to a greater number of artefact than animal stimuli (i.e., 173 and 62, respectively). If this was the case, a direct comparison between the effects of relevance for artefacts and tools should not yield any significant results. This is because the statistical power of such a comparison is not dependent on the relative number of items in the two categories. However, this direct comparison confirmed that the effect of semantic relevance was higher for artefacts than animals (see

Results and Table 2 for details). This suggests that the differential effect of relevance for artefacts and animals cannot be accounted for by a greater number of artefact than animal stimuli. An alternative explanation is that the differential results of relevance for artefacts and animals may be due to differences in the ranges of semantic relevance. As discussed in the Introduction, artefacts may have low, medium or high relevance (range: 112.29-401.13) whereas animals can only have low or medium relevance (105.44-219.03). It is therefore possible that a parametric effect of relevance is expressed across a wide range which include low, medium and high values but not when only low and medium values are available. In other words, the parametric effect of relevance might be driven by high relevance items which are available for artefacts but not for animals. This explanation is supported by the percent signal changes observed in the medial fusiform gyrus for low, medium and high relevance items, see Fig. 3. It can be seen that, in both left and right medial fusiform, percent signal change varies only minimally between low and medium relevance items, irrespective of category. Indeed, a direct statistical comparison between low and medium relevance items did not yield any significant effects even when lowering the statistical threshold to P < 0.001 (uncorrected). Furthermore, the interaction between relevance and category (i.e., low > medium relevance for artefacts vs. animals) did not reach significance using a similar statistical threshold. However, percent signal change for high relevance artefacts is much greater than for low and medium relevance items, as confirmed by a direct statistical comparison (high relevance arefects > all other conditions: left: Talairach co-ordinates z = -28; y = -52; z = -14; Z score: 5.7 P < 0.05corrected; right: Talairach co-ordinates x = 32; y = -50; z = -16; Z score: 6.0 P < 0.05 corrected). Thus, we conclude that the most likely explanation for the differential effect of relevance for artefact and animal items is that the effects are driven by high relevance items which are available as artefacts but not as animals

Table 3

Comparison between animals and artefacts, irrespective of visual relevance ("relevance unmatched") and for items with low and medium visual relevance only ("relevance matched")

Artefacts > Animals	Relevance unmatched		Relevance matched												
	Low, n	nedium an	d high rel	evance	Low a	ind medi	um relev	ance	Medi	um rel	evance		Low re	leva	nce
Anatomical location	x	у	Ζ	Z score	x	у	Ζ	Z score	x	у	Z	Z score	x y	Ζ	Z score
Left Medial Fusiform	-28	-52	-14	5.4	-28	-52	-16	3.7	-28	-52	-14	3.5	n.s.		
Left Cerebellum	-26	-36	-24	4.1		n.s.				n.s.			n.s.		
Right Medial Fusiform	32	-50	-16	5.3	30	-48	-18	3.4	32	-50	-16	3.0	n.s.		
Right Cerebellum	28	-42	-22	5.9	28	-44	-22	4.7	28	-42	-22	4.3	n.s.		

Co-ordinates [x, y, z] are reported in Talairach space. Z scores significant at P < 0.05 (corrected for multiple comparisons) are reported in bold. The remaining Z scores are significant at P < 0.001 (uncorrected). n.s. = not significant at P < 0.001 (uncorrected).



Fig. 3. The effect of artefacts > animals and its modulation by semantic relevance (thresholded at P < 0.0001 uncorrected). The bilateral medial fusiform gyri expressed greater activation for artefacts than animals. The graphs show the percent signal change for low relevance animals (Low Anim), medium relevance animals (M Anim), low relevance artefacts (Low Art), medium relevance artefacts (M Art) and high relevance artefacts (High Art).

We now move to the question of why the effect of artefacts > animals was reduced when relevance was matched for the two categories. We considered the possibility that this might be due



Fig. 4. The effect of animals > artefacts (thresholded at P < 0.0001 uncorrected). The right lateral fusiform gyri expressed greater activation for animals than artefacts. The graph shows the percent signal change for low relevance animals (Low Anim), medium relevance animals (M Anim), low relevance artefacts (Low Art), medium relevance artefacts (M Art) and high relevance artefacts (High Art).

to differences in statistical power, with the comparison including all artefacts being more sensitive than the one including low and medium relevance artefacts only. If this was the case, however, the strength of the category effects should depend on how many items are included in each category which was clearly not the case. For instance, when we compared high relevance artefacts against low and medium frequency animals, the effect of artefacts > animals became even more significant despite the number of animal stimuli (i.e., 62) being greater than the number of artefact stimuli (i.e., 23) (left: Talairach co-ordinates z = -28; y = -52; z = -14; Z score: 6.3 P < 0.05 corrected; right: Talairach co-ordinates x = 32; y = -50; z = -16; Z score: 6.4 P < 0.05 corrected). This suggests that differences in statistical power can not account for the fact that the effect of artefacts > animals was reduced when semantic relevance was matched. An alternative explanation is that differences between artefact and animal categories in the medial fusiform gyri, are driven by items with high relevance which can only be artefacts. This interpretation is supported by the percent signal changes for low, medium and high relevance items observed in the medial fusiform gyri, see Fig. 3. It can be seen that the effect of artefacts > animals is largely driven by the increased response to high relevance artefacts rather then low and medium relevance items. Thus, we conclude that the most likely explanation for why the effect of artefacts > animals was reduced when relevance was matched for the two categories, is that category effects in the medial fusiform gyri are driven by high relevance items. This suggests that the category differences typically observed in this region may be an epiphenomenon of semantic relevance rather than effects due to category per se.

It is interesting to notice that, although category effects no longer reached significance when relevance was matched, a trend could still be detected for items with medium relevance (see Results for details). One possibility is that this effect was simply due to noise, consistent with the observation that a trend could *not* be detected when the analysis was limited to low relevance items. However, it is also possible that semantic relevance does not entirely explain the category differences typically reported in the medial fusiform gyri. In other words, increased activation in the medial fusiform gyri for artefacts > animals might be a compound effect of semantic relevance and other factors. These might include some visual and semantic properties of the stimuli that were not controlled in the present study or even category per se.

A question of interest is whether semantic relevance might modulate activation in other regions of the brain that typically express category effects in functional imaging studies. For instance, when lowering the statistical threshold to P < 0.001(uncorrected), we found a trend in the right lateral fusiform gyrus which is consistent with previous studies of category effects (Chao et al., 1999; Rogers et al., 2005). However, this area did not appear to be either positively or negatively modulated by semantic relevance (P > 0.001 uncorrected) as shown in Fig. 4. It is therefore important to recognise that semantic relevance may explain increases for artefacts > animals in the medial fusiform gyrus but does not appear to drive effects in other areas that are modulated by category.

In short, the majority of previous functional imaging studies of the semantic system had looked for divisions by varying stimulus category (Ishai et al., 2000a,b; Chao et al., 1999; Damasio et al., 1996; Martin et al., 1996; Moore and Price, 1999; Mummery et al., 1996, 1998; Perani et al., 1995, 1999; Grabowski et al., 1998). These studies had led to divergent results, which may reflect differences in the types of stimuli that have been presented. To our knowledge, no attempt has been made to measure directly the effect of higher-order visual features on brain responses during picture naming. Here, we have used a quantitative index of concept activation and retrieval, namely semantic relevance. This has allowed us to demonstrate that medial fusiform activation previously reported for artefacts > animals is clearly related to the higher-order visual features, which are directly indexed by relevance. More specifically, we have shown that the medial fusiform activation is primarily driven by a subset of artefacts which have high semantic relevance. When these items are excluded, category effects are greatly diminished.

Acknowledgment

This work is funded by the Wellcome Trust.

Appendix A

	Semantic
ADTEEACTS	Terevance
ARIEFACTS	
High relevance	
BUS	401.13
MOTORCYCLE	400.15
BICYCLE	375.92
GUN	372.69
HOUSE	371.62
CHURCH	370.69
OLL	364.76

Appendix	А	(continued)
----------	---	-------------

	Semantic
ARTEFACTS	1010+41100
High relevance	
COUCH	364.54
LAMP	363.46
SAND-GLASS	362.61
SHOE	357.69
SAIL BOAT	354.53
ROCKING CHAIR	353.62
WINDMILL WELL	347.30
WELL SNOWMAN	342.07
NEST	337.61
EASEL	337.39
GLASSES	336.46
BED	333.30
HELICOPTER	328.23
CHAIN	327.89
SHIRT	327.48
Medium relevance	
SHIELD	324.38
CAR	324.31
JACKET	324.03
CANDLE	323.88
FLASK	321.38 310.30
RAFT	319.30
PIANO	313.61
TELEPHONE	310.61
DRESS	307.83
TRUCK	307.65
SWEATER	306.25
COAT	305.01
TELEVISION	304.86
STOVE	303.00
TORCH	302.40
BELL	301.94
CANOPY	299.48
CLOCK	293.77
PULPH KETTI E	291.57
FENCE	290.05
STOOL	285.86
MICROSCOPE	284.73
CHAIR	280.29
FORK	279.31
LIGHTHOUSE	279.13
MASK	278.94
AIRPLANE	278.47
NEEDLE	277.85
BELT	277.63
POCKETBOOK	277.00
ACCORDION	2/5.4/
BOOT	270.51
DESK	200.92
AXE	266.25
TABLE	264.93
SCALE	263.85
BRUSH	263.72
IGLOO	259.77
RECORD PLAYER	258.46
COMB	257.07

(continued on next page)

Appendix A (continued)

Appendix A (continued)

	Semantic		Semantic
	relevance		relevance
ARTEFACTS		ARTEFACTS	
Medium relevance	254.02	Medium relevance	100.02
LUCK	254.92	BARKEL	190.03
FLAG	253.19	PEN	188.46
PENCIL	255.15	HANGER	185.68
ADMOUD	252.79	Low relevance	165.06
LIGHT BULB	252.54	PARACHUTE	182 33
SUITCASE	252.00	CROWN	181.85
NUT	251.40	PIPE	181.00
TELESCOPE	250.53	KEY	180.13
ROLLER SKATE	248.62	SPOON	179.51
ASHTRAY	245.38	BUTTON	178.44
IRON	243.68	ANCHOR	176.78
FOOTBALL HELMET	243.61	FRYING PAN	175.82
BABY CARRIAGE	243.36	CIGAR	173.08
CIGARETTE	242.53	BALLOON	172.70
SPINNING WHEEL	241.58	SIEVE	172.54
TRAIN	240.58	CAP	171.92
DIRIGIBLE	240.15	HARP	171.78
PLUG	236.03	BASKEI	168.66
SUISSORS	235.76	VASE DVD A MID	166.40
LIGHT SWITCH	235.03	SI ED	165.43
VIOLIN	234.03	POT	165.45
DOORKNOR	232.71	KITE	164.75
KNIFF	231.49	SPOOL OF THREAD	164.52
WINEGLASS	228.80	PAGODA	164.40
IRONING BOARD	227.87	FRENCH HORN	161.69
CANNON	225.66	RING	160.37
OIL CRUET	223.39	TOP	154.83
REFRIGERATOR	223.28	UMBRELLA	152.84
ANVIL	223.06	PAINTBRUSH	152.46
TENNIS RACKET	221.18	GLOVE	152.26
TOOTHBRUSH	220.57	ROLLING PIN	150.52
TURBAN	220.30	SCREWDRIVER	149.21
TOASTER	219.89	BARN	147.65
RULER	219.87	HAMMOUK	146.82
BOOK	218.38	BROOM	146.09
WAGON	218.25	GLASS	130.60
SAW EQOTRALI	217.72	CLIP	137.68
RASEBALL BAT	216.54	DRUM	136.52
PANTS	213.78	BOW	125.78
WATERING CAN	213.70	THIMBLE	112.29
TUB	212.30		
BALL	212.23	ANIMALS	
GUITAR	211.29	Medium relevance	
CLOTHESPIN	210.19	ELEPHANT	325.43
TRAFFIC LIGHTS	209.79	GOAT	278.90
HAT	209.63	LOBSTER	258.91
DOOR	207.48	KANGAROO	255.73
WHISTLE	206.86	FISH	252.75
SOCK	205.31	ROOSTER	245.30
BIGA	204.21	HEN	244.22
NAIL	203.98		239.99
SWORD	202.60	CAWEL	237.01
UAKE	202.22	SNAII	255.45
	201.22	PEACOCK	252.01
KADAK DI IERS	200.34	DONKEY	227.55
GARBAGE CAN	190.00	SNAKE	215.08
SALT SHAKER	197.61	EAGLE	214.25
SALI SHAKEK	192.01		217.23

Appendix A (continued)

	Semantic relevance
ANIMALS	
Medium relevance	
HIPPOPOTAMUS	212.46
SEAL	208.84
HORSE	205.48
RHINOCEROS	205.45
BUTTERFLY	198.93
MOUSE	196.45
SEA HORSE	195.76
BIRD	190.89
DEER	190.48
GIRAFFE	189.82
TIGER	188.53
GRASSHOPPER	187.72
TURKEY	186.73
LION	185.76
ZEBRA	184.95
PARROT	184.66
Low relevance	
TURTLE	183.43
DUCK	183.08
LEOPARD	176.35
OSTRICH	175.91
TOUCAN	175.59
BEETLE	175.18
SWAN	173.35
OWL	170.27
SHEEP	170.00
WHALE	168.55
BEE	168.33
ANT	166.20
PENGUIN	161.22
CAT	160.37
GORILLA	155.07
RACCOON	152.52
FLY	150.33
CATERPILLAR	150.23
BEAR	147.91
FOX	143.37
SKUNK	140.51
SQUIRREL	140.10
PIG	140.10
WOLF	139.40
DACHSHUND	139.40
ALLIGATOR	138.10
MONKEY	137.12
PANDA	134.84
RABBIT	130.43
FROG	127.71
SPIDER	105.44

References

- Allport, D.A., 1985. Distributed memory, modular subsystems and dysphasia. In: Newman, S., Epstein, R. (Eds.), Current Perspectives in Dysphasia. Churchill Livingstone, Edinburgh, pp. 32–60.
- Ashcraft, M.H., 1978. Property dominance and typicality effects in feature statement verification. J. Verbal Learn. Verbal Behav. 17, 155–164.
- Bookheimer, S.Y., Zeffiro, T.A., Blaxton, T., Gailard, W., Theodore, W., 1995. Regional cerebral blood flow during object naming and word reading. Hum. Brain Mapp. 3, 93–106.
- Cavanagh, P., Arguin, M., Treisman, A., 1990. Effect of surface medium on

visual search for orientational size features. J. Exp. Psychol. Hum. Percept. Perform. 16, 479-491.

- Chao, L.L., Haxby, J.V., Martin, A., 1999. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat. Neurosci. 2, 913–919.
- Cree, G.S., McRae, K., 2003. Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello and many other such concrete nouns. J. Exp. Psychol. Gen. 132, 163–201.
- Damasio, H., Grabowski, T.J., Tranel, D., Hichwa, R.D., Damasio, A.R., 1996. A neural basis for lexical retrieval. Nature 380, 499–505.
- Dell'Acqua, R., Lotto, L., Job, R., 2000. Naming time and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. Behav. Res. Meth. Instrum. Comput. 32, 588–615.
- Devlin, J.T., Russell, R.P., Davis, M.H., Price, C.J., Moss, H.E., Fadili, J., Tyler, L.K., 2002. Is there an anatomical basis for category specificity? Semantic memory studies in PET and fMRI. Neuropsychologia 40, 54–75.
- Farah, M.J., McClelland, J.L., 1991. A computational model of semantic impairment: Modality specificity and emergent category specificity. J. Exp. Psychol. 120, 339–357.
- Friston, K.J., Ashburner, J., Frith, C.D., Poline, J.B., Heather, J.D., Frackowiak, R.S.J., 1995a. Spatial registration and normalization of images. Hum. Brain Mapp. 2, 1–25.
- Friston, K.J., Holmes, A., Worsley, K.J., Poline, J.B., Frith, C.D., Frackowiak, R.S.J., 1995b. Statistical parametric maps in functional imaging: A general linear approach. Hum. Brain Mapp. 2, 189–210.
- Garrard, P., Lambon Ralph, M.A., Hodges, J.R., Patterson, K., 2001. Prototypicality, distinctivness and intercorrelation: Analyses of the semantic attributes of living and non-living concepts. Cogn. Neuropsychol. 18, 125–174.
- Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000. The fusiform "face area" is part of a network that processes faces at the individual level. J. Cogn. Neurosci. 12, 495–504.
- Grabowski, T.J., Damasio, H., Damasio, A.R., 1998. Premotor and prefrontal correlates of category-related lexical retrieval. NeuroImage 7, 232–243.
- Hauk, O., Johnsrude, I., Pulvermuller, F., 2004. Somatotopic representation of action words in human motor and premotor cortex. Neuron 41, 301–307.
- Ishai, A., Ungerleider, L.G., Martin, A., Schouten, J.L., Haxby, J.V., 2000a. The representation of objects in the human occipital and temporal cortex. J. Cogn. Neurosci. 12 (Supplement 2), 35–51.
- Ishai, A., Ungerleider, L.G., Haxby, J.V., 2000b. Distributed neural systems for the generation of visual images. Neuron 28, 979–990.
- Kellenbach, M.L., Hovius, M., Patterson, K., 2005. A PET study of visual and semantic knowledge about objects. Cortex 41, 121–132.
- Martin, A., Chao, L.L., 2001. Semantic memory and the brain: structure and processes. Curr. Opin. Neurobiol. 11, 194–201.
- Martin, A., Haxby, J.V., Lalonde, F.M., Wiggs, C.L., Ungerleider, L.G., 1995. Discrete cortical regions associated with knowledge of color and knowledge of action. Science 270, 102–105.
- Martin, A., Wiggs, C.L., Ungerleider, L.G., Haxby, J.V., 1996. Neural correlates of category-specific knowledge. Nature 379, 649–652.
- Moore, C.J., Price, C.J., 1999. A functional neuroimaging study of the variables that generate category-specific object processing differences. Brain 122, 943–962.
- Mummery, C.J., Patterson, K., Hodges, J.R., Wise, R.J., 1996. Generating 'tiger' as an animal name or a word beginning with T: Differences in brain activation. Proc. R. Soc. London, B 263, 989–995.
- Mummery, C.J., Patterson, K., Hodges, J.R., Price, C.J., 1998. Functional neuroanatomy of the semantic system: Divisible by what? J. Cogn. Neurosci. 10, 766–777.
- Murtha, S., Chertkow, H., Beauregard, M., Evans, A., 1999. The neural substrate of picture-naming. J. Cogn. Neurosci. 11, 399–423.
- Noppeney, U., Price, C.J., 2003. Functional imaging of the semantic system: Retrieval of sensory-experienced and verbally learned knowledge. Brain Lang. 84, 120–133.

- Perani, D., Cappa, S.F., Bettinardi, V., Bressi, S., Gorno-Tempini, M., Matarrese, M., Fazio, F., 1995. Different neural systems for the recognition of animals and man-made tools. NeuroReport 6, 1637–1641.
- Perani, D., Cappa, S.F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M., Fazio, F., 1999. The neural correlates of verbs and nouns processing: a PET study. Brain 122, 2337–2344.
- Price, C.J., Devlin, J.T., Moore, C.J., Morton, C., Laird, A.R., 2005. Metaanalyses of object naming: Effect of baseline. Hum. Brain Mapp. 25, 70–82.
- Rogers, T.T., Hocking, J., Mechelli, A., Patterson, K., Price, C.J., 2005. Fusiform activation to animal is driven by the process, not the stimulus. J. Cogn. Neurosci. 17, 434–445.
- Saffran, E.M., 2000. The organization of semantic memory: In support of a distributed model. Brain Lang. 71, 204–212.
- Sartori, G., Lombardi, L., 2004. Semantic relevance and semantic disorders. J. Cogn. Neurosci. 16, 439–452.
- Sartori, G., Lombardi, L., Mattiuzzi, L., 2005. Semantic relevance better

predicts normal and abnormal concept retrieval. Neuropsychologia 43, 754-777.

- Snodgrass, J.G., Vanderwart, M., 1980. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. J. Exp. Psychol. Hum. Learn. Mem. 6, 174–215.
- Thompson-Schill, S.L., 2003. Neuroimaging studies of semantic memory: Inferring "how" from "where". Neuropsychologia 41, 280–292.
- Thompson-Schill, S.L., Aguirre, G.K., D'Esposito, M., Farah, M.J., 1999. A neural basis for category and modality specificity of semantic knowledge. Neuropsychologia 37, 671–676.
- Tyler, L.K., Moss, H.E., Durrant-Peatfield, M., Levy, J., 2000. Conceptual structure and the structure of the concepts: a distributed account of category specific deficits. Brain Lang. 75, 195–231.
- Warrington, E.K., McCarthy, R., 1987. Categories of knowledge: Further fractionations and an attempted integration. Brain 110, 1273–1296.
- Warrington, E.K., Shallice, T., 1984. Category-specific semantic impairments. Brain 107, 829–853.