

## Right fusiform response patterns reflect visual object identity rather than semantic similarity

Rose Bruffaerts<sup>1,2</sup>, Patrick Dupont<sup>1</sup>, Sophie De Grauwe<sup>1,3</sup>, Ronald Peeters<sup>4</sup>, Simon De Deyne<sup>5</sup>, Gerit Storms<sup>5</sup>, Rik Vandenberghe<sup>1,2</sup>

<sup>1</sup>Laboratory for Cognitive Neurology, Department of Neurosciences, KU Leuven, Belgium; <sup>2</sup>Neurology Department, University Hospitals Leuven, Leuven, Belgium; <sup>3</sup>Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behavior, Nijmegen, The Netherlands; <sup>4</sup>Radiology Department, University Hospitals Leuven, Leuven, Belgium; <sup>5</sup>Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium

*Corresponding author:* Rik Vandenberghe, M.D., Ph.D., Neurology Department, University Hospitals Leuven, Herestraat 49, 3000 Leuven, Belgium, e-mail: rik.vandenberghe@uz.kuleuven.ac.be, phone: ++ 32 (0)16 344280, fax: ++ 32 (0)16 344285

*This is a pre-copy-editing, author-produced PDF of an article accepted for publication in NeuroImage following peer review. The definitive publisher-authenticated version (Bruffaerts, R., Dupont, P., De Grauwe, S., Peeters, R., De Deyne, S., Storms, G., and Vandenberghe, R. (2013). Right fusiform response patterns reflect visual object identity rather than semantic similarity, 83: 8797) is available online at: <http://dx.doi.org/10.1016/j.neuroimage.2013.05.128>.*

## **Abstract**

We previously reported the neuropsychological consequences of a lesion confined to the middle and posterior part of the right fusiform gyrus (case JA) causing a partial loss of knowledge of visual attributes of concrete entities in the absence of category-selectivity (animate versus inanimate). We interpreted this in the context of a two-step model that distinguishes structural description knowledge from associative-semantic processing and implicated the lesioned area in the former process. To test this hypothesis in the intact brain, multi-voxel pattern analysis was used in a series of event-related fMRI studies in a total of 46 healthy subjects. We predicted that activity patterns in this region would be determined by the identity of rather than the conceptual similarity between concrete entities. In a prior behavioral experiment features were generated for each entity by more than 1000 subjects. Based on a hierarchical clustering analysis the entities were organized into 3 semantic clusters (musical instruments, vehicles, tools). Entities were presented as words or pictures. With foveal presentation of pictures, cosine similarity between fMRI response patterns in right fusiform cortex appeared to reflect both the identity of and the semantic similarity between the entities. No such effects were found for words in this region. The effect of object identity was invariant for location, scaling, orientation axis and color (grayscale versus color). It also persisted for different exemplars referring to a same concrete entity. The apparent semantic similarity effect however was not invariant. This study provides further support for a neurobiological distinction between structural description knowledge and processing of semantic relationships and confirms the role of right mid-posterior fusiform cortex in the former process, in accordance with previous lesion evidence.

## 1. Introduction

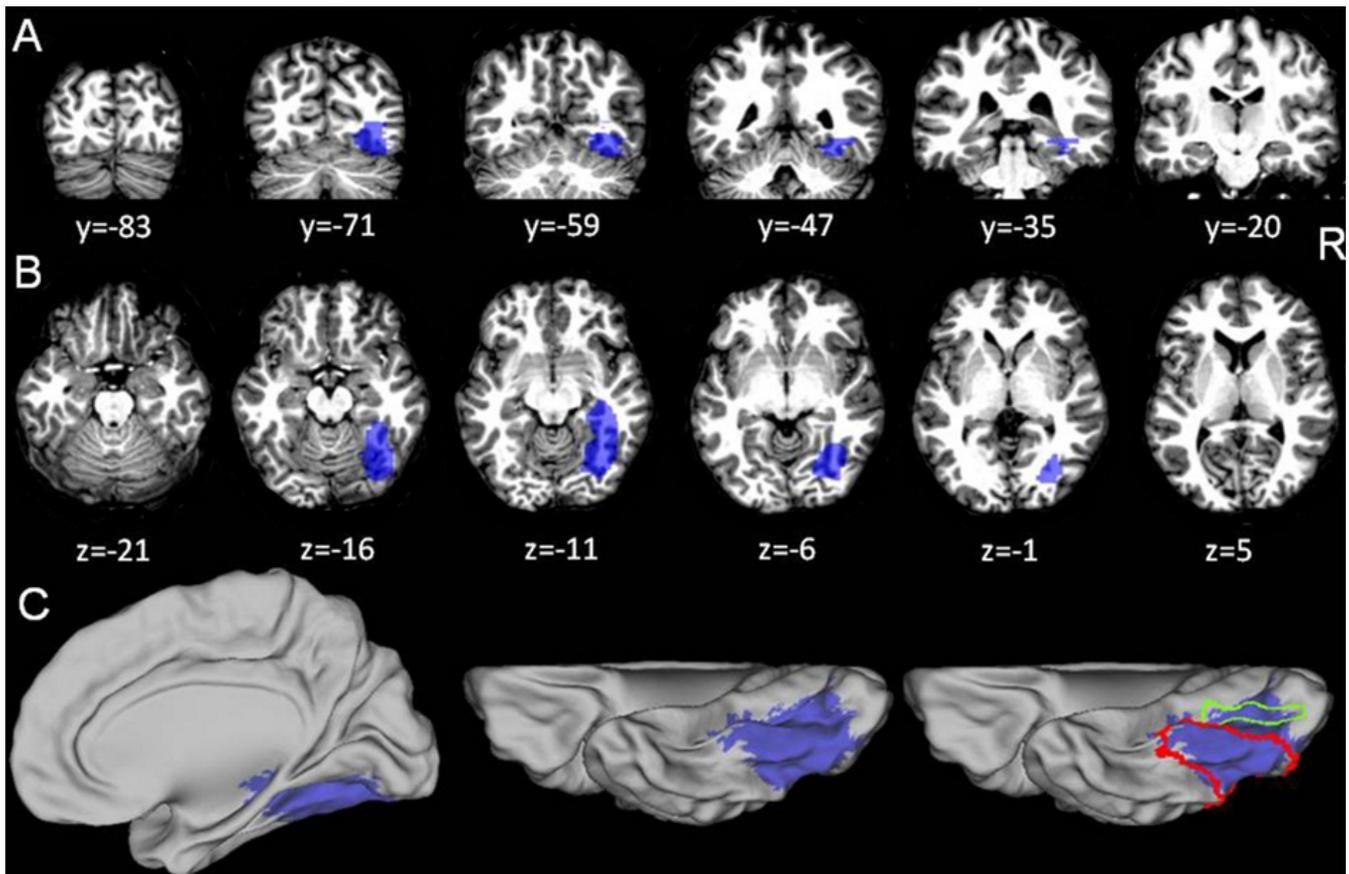
Elucidating the functional organisation of occipitotemporal cortex is crucial for understanding how visuoperceptual input is relayed to the semantic network (Farah, 2004; Humphreys and Riddoch, 2003; Logothetis and Sheinberg, 1996; Marr, 1982; Martin et al., 1999; Rogers and McClelland, 2004). Neuropsychology (Farah, 2004; Humphreys and Riddoch, 2003) and non-invasive mapping of the intact brain (e.g. Lerner et al. (2001); Price and Devlin (2003); Devlin et al. (2005); Bright et al. (2005); Huth et al. (2012); Peelen and Caramazza (2012)) continue to play a prominent role in this endeavour, among other methods (Marr, 1982; Biederman, 1987; Logothetis and Sheinberg, 1996; Nobre et al., 1994; Tyler and Moss, 2001; Rogers and McClelland, 2004; Liu et al., 2009; Chan et al., 2011; DiCarlo et al., 2012). Theories differ in how basic-level identification of concrete entities relates to memory for semantic relationships between these entities. In healthy subjects object identification almost automatically provides access to the meaning of the object and activates to varying degrees the connections of the concrete entity within the semantic 'web'. Computational models propose that a concrete entity (such as e.g. 'cat') belongs to a semantic web consisting of a myriad of nodes (e.g. other concrete entities) which are connected with variable connection strengths (Dell, 1986; Rogers and McClelland, 2004). In the current paper, we will use object identification to refer to the 'nodal' level and semantic processing to the 'network' level, although we realize that the 'nodes' themselves, that correspond to entities at the basic level, are also constituted by combinations of different features (object parts, shape ...) (Dell, 1986). From this it should be clear that, neurobiologically, to disentangle basic-level object identification from the automatic retrieval of semantic relationships and associations constitutes a difficult problem and that it requires converging evidence from multiple experimental sources.

The distinction between object identification and associative-semantic processing may even be artificial: The 'embodied cognition' concept emphasizes the intimate relationship between visuoperceptual identification and categorical processing (Barsalou, 2008; Martin et al., 1999). According to such models, category-specificity arises from the structure of the feature space as represented in occipitotemporal cortex (Martin and Chao, 2001). Functional imaging studies in healthy controls have indicated that ventral occipitotemporal cortex may be best viewed as a lumpy feature-space, representing stored information about features of object form shared by members of a category (Martin and Chao, 2001).

According to an alternative theory, the Hierarchical Interactive Theory (Forde et al., 1997; Humphreys and Riddoch, 2003), the structural description system mediates object identification and is neurobiologically distinct from the system mediating associative-semantic processing. According to this model, object recognition consists of at least two neurobiologically separable steps: one step permits the matching of the encoded representation of an object to the corresponding stored structural description, the other maps structural knowledge onto the 'semantic representations' (Hillis and Caramazza, 1995; Rumiati and Humphreys, 1997). Structural descriptions code for the long-term memory of perceptual features that are necessary for the identification of familiar objects, in particular shape and object parts (Forde et al., 1997; Humphreys and Riddoch, 2003). One of the principal sources of empirical evidence for these fundamental questions comes from focal lesions of the occipitotemporal processing stream in humans.

Lesion case studies have provided evidence for a role of right posterior/middle fusiform cortex in invariant visuoperceptual identification (Konen et al., 2011) as well as in mnemonic retrieval of visual features of concrete entities (Vandenbulcke et al., 2006). Case JA had a lesion confined to right posterior and middle fusiform gyrus (Fig. 1) (Vandenbulcke et al., 2006). JA was impaired on the object decision task, a task that requires subjects to discriminate drawings of real-life objects from chimaeric objects and that is classically regarded as diagnostic for a structural description deficit (Riddoch and Humphreys, 1993). Perceptual identification speed was decreased in JA (Vandenbulcke et al., 2006) while copying was preserved. Case JA was impaired on tasks probing explicit

retrieval of visuoperceptual features through both verbal (feature generation, forced-choice naming to definition) and nonverbal means (drawing from memory). Despite a selective impairment of retrieval of visual features, JA did not demonstrate a semantic category effect, biological versus non-biological (Vandenbulcke et al., 2006) and was able to name pictures of objects from different categories, although JA needed longer stimulus presentation durations than controls. We interpreted the deficit in JA as a selective loss of knowledge of visual features of real-life concrete entities and situated the deficit at the structural description processing level. According to these lesion data, this structural description knowledge can be retrieved explicitly through verbal or non-verbal means (Vandenbulcke et al., 2006). A second relevant case, SM, had a traumatic lesion similar in location to JA's lesion. SM displayed integrative object agnosia (Behrmann and Kimchi, 2003a,b; Behrmann and Williams, 2007), characterized by impaired spatial integration of object parts (Behrmann et al., 2006). Recognition of a visual stimulus that represents a real-life entity entails identification of the stimulus as an instantiation of that entity in a way that tolerates variations in viewpoint, stimulus position, scaling etc. (Rust and Dicarlo, 2010; DiCarlo et al., 2012). Viewpoint invariance, a prerequisite for the processing of structural descriptions, was bilaterally reduced in the Lateral Occipital Complex (LO) in case SM (Konen et al., 2011). A third case, DHY (Hillis and Caramazza, 1995), who had optic aphasia, fundamentally differed from JA and SM. In DHY, access to structural description knowledge was preserved while processing of fine-grained semantic distinctions (e.g. sorting between cats and dogs, or edible versus non-edible animals) was impaired (Hillis and Caramazza, 1995). Taken together, these cases provide evidence for a neurobiological distinction between object identification and processing of fine-grained semantic relationships between objects.



**Figure 1** – Lesion map of case JA (blue). **A.** Coronal sections. **B.** Transverse sections. **C.** Projections onto a surface rendering of the brain (Caret v5.65 (Van Essen, 2005)). Green outline: area V4v (green, extracted from Caret v5.65 (Van Essen, 2005)). Red outline: area LO, defined by contrasting the picture stimuli of Kourtzi and Kanwisher (2001) with the scrambled pictures from that study (Vandenbulcke et al., 2006).

In the current fMRI study, we tested the hypothesis that the region that was lesioned in JA, right mid-posterior fusiform cortex (Fig. 1), is involved in structural description processing which is necessary for invariant basic level object identification and that this process can be neurobiologically dissociated from the processing of semantic similarity. We used multi-voxel pattern analysis (MVPA) to determine whether responses in right posterior/middle fusiform cortex reflect object identity versus semantic similarity in cognitively intact subjects. An object identity effect was defined as a higher cosine similarity between activity patterns elicited by replications of a same entity despite variations in location, scaling, color, orientation and exemplar. A semantic similarity effect was defined as a higher cosine similarity between response patterns elicited by entities belonging to a same semantic cluster that persists despite variations in location, scaling, color, orientation and exemplar. This invariance criterion is essential to discriminate hidden sensory confounds from true effects of object identity or category (DiCarlo et al., 2012).

In the main experiment we evaluated the effect of object identity and semantic similarity with foveal presentation of words and pictures, in the invariance experiments we examined how variations in parafoveal location, size, orientation and color of pictures affected the effects obtained, and in the exemplars experiment we assessed the effect of object kind identity while varying the exemplar representing a concrete entity.

## 2. Materials and Methods

### 2.1. Participants

Twenty subjects (12 women, 8 men, between 18 and 25 years old) participated in the main experiment. Eight (6 women, 2 men, between 20 and 25 years old), 10 (6 women, 4 men, between 19 and 26 years old) and 8 volunteers (6 women, 2 men, between 18 and 26 years old), respectively, participated in the two invariance experiments and the exemplars experiment. All subjects were native Dutch speakers and strictly right-handed (Oldfield, 1971). The volunteers were free of psychotropic or vasoactive medication and had no neurological or psychiatric history. All participants gave written informed consent in accordance with the Declaration of Helsinki. The Ethics Committee of the University Hospitals Leuven approved the experimental protocol.

### 2.2. Volumes of interest

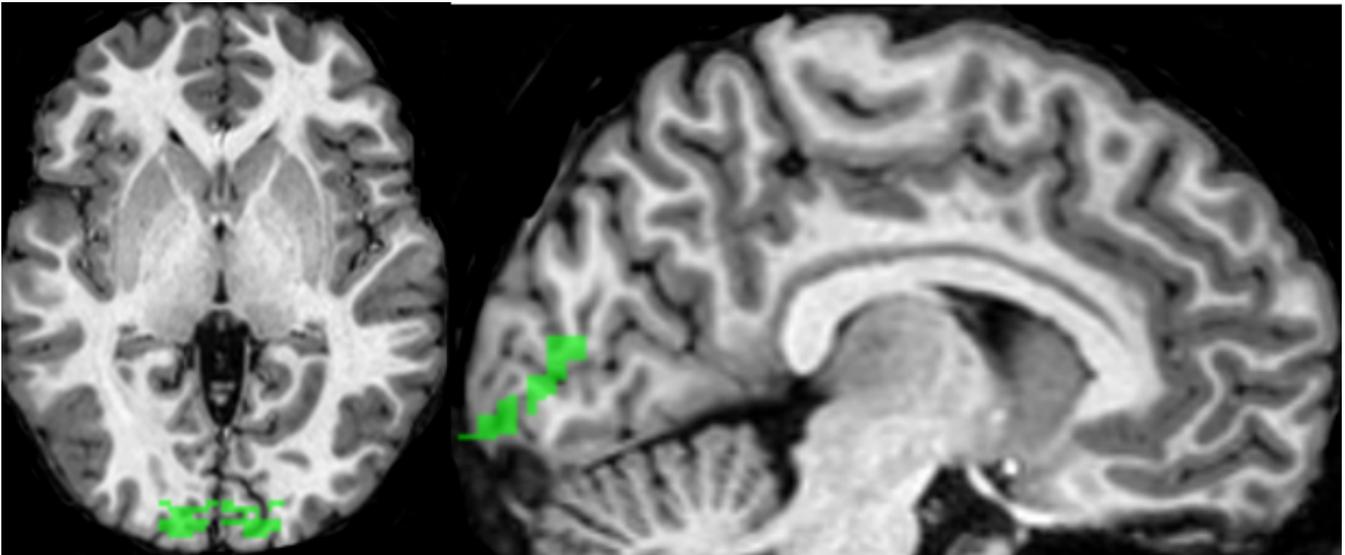
Our main volume of interest (VOI) was directly imported from patient JA's lesion map (Vandenbulcke et al., 2006), which forms the basis for the current study (Fig. 1). The lesion corresponded to right middle and posterior fusiform gyrus (1785 voxels of  $2 \times 2 \times 2 \text{ mm}^3$ ; Montreal Neurological Institute (MNI) coordinates of the boundaries:  $x = 20$  to  $50 \text{ mm}$ ,  $y = -81$  to  $-21 \text{ mm}$ ,  $z = -20$  to  $4 \text{ mm}$ ). The lesion volume will be denoted as 'the JA lesion VOI'. This lesion overlaps to a significant degree with the ventral portion of LO as defined by a functional localizer based on the contrast between pictures of real objects and scrambled versions of these pictures (Kourtzi and Kanwisher, 2001; Vandenbulcke et al., 2006) (Fig. 1).

We also evaluated the left-sided mirror region and included as a control VOI area V1 (784 voxels), as extracted from Wake Forest University (WFU) PickAtlas version 2.4 (Maldjian et al. (2003)) (Fig. 2).

### 2.3. Main experiment

#### 2.3.1. Stimuli

Our selection and characterization of entities to be used for fMRI was based on an existing behavioral dataset (De Deyne et al., 2008). This dataset was assembled as follows: 1003 college students performed a feature generation task with 229 concrete inanimate entities (De Deyne et al., 2008). Each student was presented with 10 stimulus words, and asked to write down 10 features for

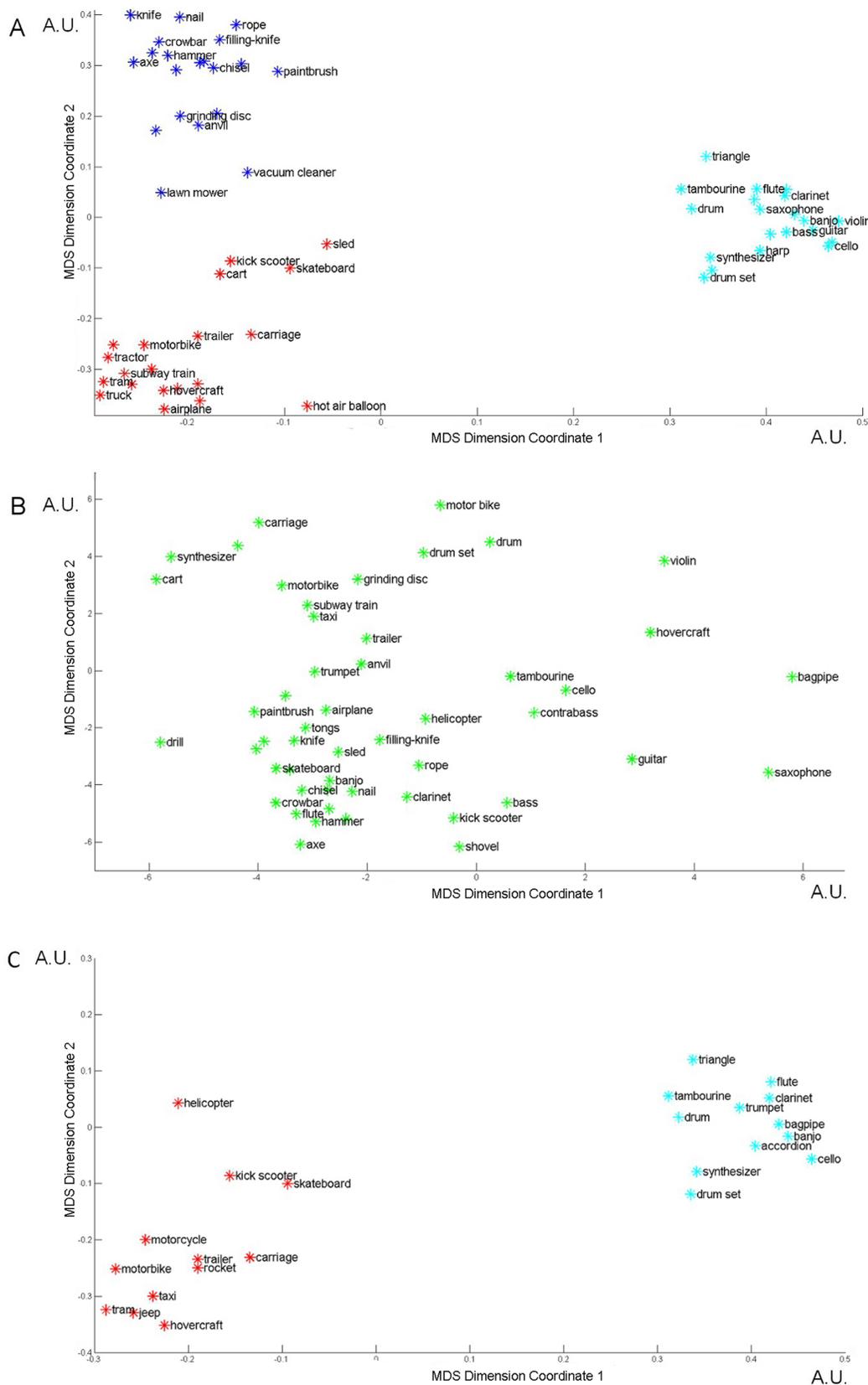


**Figure 2** – Primary visual control region, extracted from Wake Forest University (WFU) PickAtlas version 2.4 (Maldjian et al. (2003))

each word, emphasizing that different types of features had to be generated (e.g. perceptual, functional). A minimum of 180 features was gathered for each stimulus word. Next, the applicability of the most frequently generated features was scored for each word by 5 different subjects, who were instructed to judge for each entity-feature pair whether or not the feature characterized the entity (feature applicability judgment task) (De Deyne et al., 2008). This matrix, with rows corresponding to objects and columns to semantic features (concept-feature matrix), was used to derive a similarity matrix by computing the cosine similarities between each row. Note that such a similarity matrix reflects all kinds of features including encyclopedic or sensory ones. Data reduction of the similarity matrix for visualization in two dimensions was performed and a bottom-up hierarchical clustering method was applied, by means of multidimensional scaling (MDS) and Ward's method (for minimizing the within-cluster sum of squares), as implemented in Matlab 2009b, Statistical Toolbox.

For the current study, 20 stimuli were selected from each of three semantic clusters (musical instruments, vehicles and tools) (Fig. 3A). The stimuli from the 3 semantic clusters were matched for word frequency (Baayen et al. (1993), Dutch version), word generation frequency, age of acquisition, familiarity, imageability and word length (Table 1) (De Deyne et al., 2008). Word length was between 3 and 11 characters. For each noun, a prototypical color photo was selected. Familiarity and visual complexity of these pictures were rated by 38 and 33 volunteers, respectively, using a seven-point (1-7) Likert-type scale. According to these ratings, the 3 semantic clusters were matched for picture familiarity ( $P = 0.216$ ) (Table 1). Picture complexity was lower for tools than for musical instruments and vehicles ( $P < 0.001$ ) (Table 1).

In order to evaluate how perceptual similarity related to semantic similarity, we calculated Euclidean distances between the physical characteristics of the grayscale versions of the pictures that were used for fMRI. First, we applied Gaussian filtering to these pictures (full width at half maximum: 35 x 35 pixels, picture size: 600 x 600 pixels, pixel values between 0-255 (8 bit)) (Kayaert et al. (2003)). Next, the Euclidean distance was calculated between each pair of pictures, yielding a perceptual dissimilarity matrix. To visualize the dissimilarity matrix, we reduced it to two dimensions using MDS. The clustering of the multidimensionally scaled data based on visuoperceptual characteristics was substantially different from that based on the concept-feature similarity matrix (Fig. 3B).



**Figure 3** – Semantic and perceptual similarity between the entities: **A**. Main experiment: MDS of the entities, based on concept-feature similarity matrix (De Deyne et al., 2008). The 3 semantic clusters correspond to musical instruments (cyan), vehicles (red) and tools (blue). To avoid overlap between the printed names we omitted some of the names when datapoints were too close. **B**. Main experiment: MDS of the pictures, based on the perceptual dissimilarity matrix (see Materials and Methods). **C**. Invariance and exemplar experiments: MDS of the entities, based on the concept-feature similarity matrix (De Deyne et al., 2008). Musical instruments: cyan, vehicles: red. The x and y values are in arbitrary units.

Main experiment							
	Musical Instruments		Vehicles		Tools		<i>P</i>
	mean	S.D.	mean	S.D.	mean	S.D.	
<b>Words</b>							
Word frequency	1.58	(0.68)	1.98	(0.90)	1.74	(0.92)	0.162
Generation frequency	32.2	(29.54)	17.9	(2.47)	23.4	(31.26)	0.157
Age of acquisition	8.19	(1.45)	7.65	(1.98)	7.53	(2.52)	0.376
Familiarity	2.84	(0.43)	2.88	(0.47)	3.14	(0.69)	0.257
Imageability	5.71	(0.60)	5.99	(0.60)	5.60	(0.85)	0.188
Word length	7.40	(2.06)	6.80	(2.76)	7.45	(2.91)	0.623
<b>Pictures</b>							
Familiarity	3.79	(0.69)	4.00	(0.92)	4.24	(0.99)	0.216
Complexity	3.74	(0.86)	3.87	(0.85)	2.61	(0.94)	<0.001

**Table 1 – Lexical norms and picture properties of the main experiment (mean and standard deviation):**

The reported frequency values are logarithmic function values of the lemma counts from the Dutch version of the CELEX database (Baayen et al. (1993)). Generation frequency measures how often a stimulus was produced in response to the category label, taking into account the order of generated exemplars. For the determination of the age of acquisition, subjects were asked to estimate the age at which they had first learned the stimulus word. Familiarity, complexity and imageability ratings were scored on seven-point Likert-type scales (De Deyne et al., 2008). *P* values are based on a Kruskal-Wallis rank sum test (statistics performed in R, version 2.12.1).

### 2.3.2. Experimental design

In the main experiment, each stimulus was displayed centrally on the screen for 500 msec. Letter size was 1.2 visual degrees, picture size 6.9 x 6.9 visual degrees. The stimulus was followed by a central fixation point for 10 seconds. A long interstimulus interval was chosen to minimize the need for deconvolution of the haemodynamical response to each stimulus. Each subject underwent a session of 8 runs containing 30 trials each. Each entity was presented once per two runs and in total, each of the 60 entities was presented twice as a word and twice as a picture. The order between the 6 trial types (musical instruments, vehicles and tools presented as words or pictures) was pseudorandomized using OptimizeDesign 10 (Wager and Nichols (2003)). During stimulus presentation, subjects were asked to fixate and indicate by button press whether the stimulus was presented as a word or as a picture to minimize conceptual processing.

Stimulus presentation was controlled by a PC running Presentation 14.8 (NeuroBehavioural Systems, Albany, CA, USA). Subjects viewed the stimuli on a mirror in front of them using a Barco 6400i projector at a frequency of 60 Hertz and a resolution of 1024 x 768 pixels.

### 2.4. Invariance experiments

We examined whether effects of object identity or semantic similarity observed in the main experiment were invariant for location. Invariance for location is a key criterion for true object identification/categorization and in the absence of any proof of invariance, visual sensory confounds should be considered rather than a role in true object identification/categorization (DiCarlo et al. (2012)).

For both invariance experiments, picture stimuli (size = 5.0 x 5.0 visual degrees) consisted of 12 musical instruments and 12 vehicles taken from the stimulus set used in the main experiment (Fig. 3C). Pictures were presented parafoveally to investigate location invariance.

#### 2.4.1. First invariance experiment

Pictures were presented at 2.5 degrees eccentricity in one of 4 possible quadrants centered on the diagonal. They were presented either in the original orientation or flipped along their vertical axis.

Location and orientation were counterbalanced over runs for each subject. Subjects were instructed to fixate the central fixation point throughout the run. Presentation duration was reduced to 150 ms to prevent saccades to the stimulus. The intertrial interval was 10.5 seconds. The 8 runs contained 24 trials each. Each of the pictures was presented once per run. The order between the 24 entities was pseudorandomized per subject. After the end of the MRI experiment, subjects named the parafoveal pictures. Stimuli not recognized by a subject were excluded on a subject-specific basis.

#### *2.4.2. Second invariance experiment*

In the second invariance experiment we repeated invariance testing under free viewing conditions, and additionally varied size-on-the-screen and color (prototypical color versus grayscale). A same picture could be shown in two different sizes (picture size: 5.7 x 5.7 or 6.9 x 6.9 visual degrees), in either color or grayscale, and in either the original orientation or flipped along its vertical axis. Stimulus presentation was pseudorandomized per subject, with strict counterbalancing of the location, size-on-the-screen and color variables over runs. Orientation was randomized per run. Each picture was presented for 500 ms at 4.2 degrees eccentricity centered on the diagonal in one of 4 possible quadrants. The intertrial interval was 10.5 seconds. Subjects were instructed to make a saccade from the fixation point to the picture and return to the fixation point after picture offset. Each of the 8 runs contained 48 trials, with 2 presentations of each entity.

#### *2.5. Exemplars experiment*

The same 24 entities were used as in the invariance experiments. Instead of replicating a single picture per entity, we used 8 different exemplars per entity, so that a same picture was never replicated. The 8 exemplars included the picture shown in the invariance experiments plus 7 new color photographs on a white background of different exemplars of the same basic-level entity (<http://www.google.com/imghp>). All pictures were presented foveally (4.1 x 4.1 visual degrees), with one presentation of a same entity in each of the 8 runs, but in all other respects stimulus presentation was identical to that in the second invariance experiment. Afterwards, participants named all exemplars to ensure correct identification.

To estimate low-level perceptual similarities between the picture stimuli, Euclidean distances were calculated in the same manner as in the main experiment (Kayaert et al., 2003). Average Euclidean distance between exemplars of a same entity was lower than between different entities ( $P < .001$ ).

#### *2.6. MRI Acquisition*

Structural and functional images were acquired on a 3 T Philips Intera system (Best, The Netherlands) equipped with an 8-channel head volume coil. Structural imaging sequences consisted of a T1-weighted 3D turbo-field-echo sequence (repetition time = 9.6 msec, echo time = 4.6 msec, in-plane resolution = .97 mm, slice thickness = 1.2 mm). Functional images were obtained using T2\* echoplanar images comprising 29 transverse slices oriented in parallel with the anterior-posterior axis of the temporal lobe (repetition time = 2 s, echo time = 30 msec, voxel size 1.64 x 1.64 x 2.19 mm<sup>3</sup>, slice thickness = 2 mm, Sensitivity Encoding (SENSE) factor = 2), with the field of view (FOV) (210 x 210 x 64 mm<sup>3</sup>) centered on the occipitotemporal cortex. Each run was preceded by 4 dummy scans to allow for saturation of the BOLD signal. In all experiments eye position was monitored using an Applied Science Laboratory infrared system (ASL 5000/LRO system).

#### *2.7. Image processing*

Preprocessing of the fMRI data (spatial realignment, slice time realignment, coregistration and normalization with a voxel size of 2 x 2 x 2 mm<sup>3</sup>) was performed using Statistical Parametric Mapping 2008 (SPM8) (Wellcome Trust Centre for Neuroimaging, London, UK). All images were smoothed using a 5 x 5 x 7 mm<sup>3</sup> Gaussian kernel (Op de Beeck (2010)). From these images, the

fMRI response pattern was derived by calculating the area under the curve of the BOLD response between 2 and 8 seconds after trial onset. This procedure resulted in an activation map for each stimulus presentation, containing the fMRI response pattern.

## 2.8. Multivariate analysis

For each trial, a vector was constructed in a higher-dimensional space, with the activity levels in each voxel as the elements of the vector and the dimensionality of this vector equal to the number of voxels examined. The cosine similarity between 2 vectors is the cosine of the angle formed by the vectors. When cosine similarity equals 1, the patterns are identical, save a possible scaling factor (Mitchell et al., 2008). Per subject, the cosine similarity of the vectors was calculated for each possible pair of trials within the a priori defined VOIs. The matrix containing the pairwise cosine similarity between every two trials is the similarity matrix. Cosine similarities were always determined within modality (words or pictures).

### 2.8.1. Main experiment

In the main experiment each word and each picture was replicated twice. We determined the cosine similarities between replications of a same word or a same picture and the cosine similarities between pairs belonging to a same semantic cluster.

For each word or picture replication pair, we determined the cosine similarity between the fMRI response patterns in each subject. We then averaged the cosine similarities over word pairs within each subject and did the same over the picture pairs. The average of these values across the entire group of subjects will be called the average cosine similarity (ACS) for word or picture replications. To evaluate whether the cosine similarity for word or picture replications differed from chance, we compared these cosine similarities to those obtained based on random permutation labelling, with  $10^4$  random permutations. Random permutation labelling is a non-parametric test: within each subject a random label is assigned to every trial, in a way that no label recurs. Next, cosine similarity is recalculated with these random labels for each subject. These cosine similarities obtained with random labels are then compared to the cosine similarities obtained for the replication pairs based on the true labelling. We used a one-tailed statistical threshold of  $P < 0.05$ .

Secondly, for each possible pair of entities within each semantic cluster, we determined the cosine similarity between the fMRI response patterns in each subject. We then averaged the cosine similarities of these pairs within each subject. The average of these values across the entire group of subjects will be called the average cosine similarity (ACS) within semantic clusters. This was calculated separately for words and pictures. To evaluate whether the cosine similarity within semantic clusters differed from chance, we compared these cosine similarities to those obtained based on random permutation labelling, in the same manner as outlined above.

Finally, we calculated the cosine similarity for pairs of entities coming from two different semantic clusters, using a same statistical approach. The purpose of this analysis was mainly to serve as a control to evaluate the specificity of the results obtained for within-cluster pairs.

### 2.8.2. Invariance experiments

In the invariance experiments, each of 24 pictures was shown 8 times, twice per quadrant. The 24 pictures came from two semantic clusters. We determined the cosine similarities between the 8 replications of a same picture, secondly the cosine similarities between all possible pairs belonging to a same semantic cluster, and finally between all trials sharing a same visual quadrant of stimulation. The mathematical procedure was identical to that outlined above.

For replications of a same picture, cosine similarity was calculated for each possible pair of identical pictures, regardless of stimulus location, in each subject. These values were then averaged within each subject. These cosine similarities obtained across the entire group of subjects were then compared to the probability distributions obtained with  $10^4$  random permutation labellings.

For testing the cosine similarity within semantic clusters, cosine similarity was calculated for each possible pair belonging to a same semantic cluster, regardless of stimulus location, in each subject. These values were then averaged over all within-cluster pairs per subject. These cosine similarities obtained across the entire group were then compared to the probability distributions obtained with  $10^4$  random permutation labellings.

For testing the cosine similarity between trials sharing a same stimulus quadrant, cosine similarity was calculated for each possible pair of trials sharing a same quadrant, regardless of picture identity, in each subject. These values were then averaged over all same-location pairs per subject. These cosine similarities obtained across the entire group were then compared to the probability distributions obtained with  $10^4$  random permutation labellings.

### 2.8.3. Exemplars experiment

In the exemplars experiment, each entity was foveally presented 8 times, using a picture of a different exemplar of the entity per presentation. Twenty-four entities were used in total belonging to two different semantic clusters. We determined the cosine similarities between the 8 replications of a same entity as well as the cosine similarities between all possible pairs belonging to a same semantic cluster. The mathematical procedure was identical to that outlined above.

For replications of a same entity, cosine similarity was calculated for each possible pair of pictures representing a same entity in each subject. These values were then averaged over pairs representing a same entity within each subject. These cosine similarities were obtained for each subject and the average over the entire group was then compared to the probability distributions obtained with  $10^4$  random permutation labellings.

For testing the cosine similarity within semantic clusters, cosine similarity was calculated for each possible pair belonging to a same semantic cluster, in each subject. These values were then averaged over all within-cluster pairs per subject. These cosine similarities obtained across the entire group were then compared to the probability distributions obtained with  $10^4$  random permutation labellings.

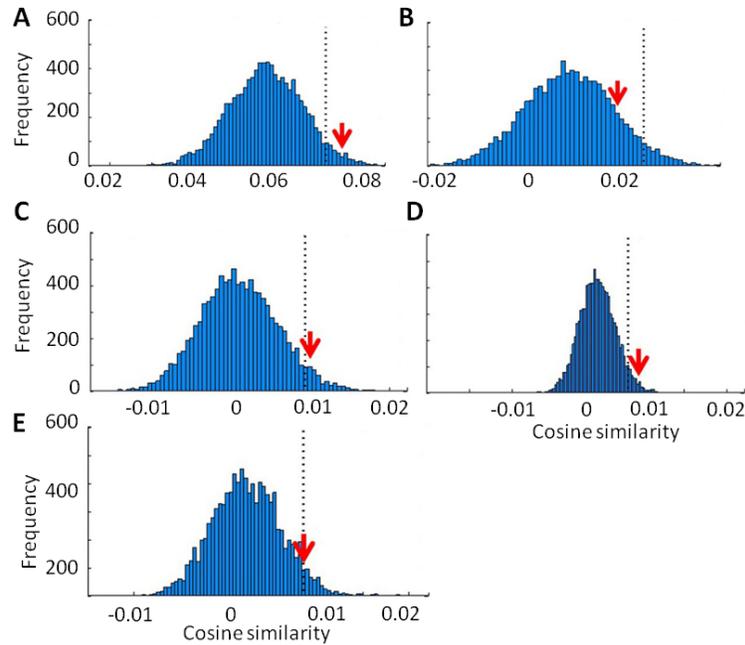
## 3. Results

### 3.1. Main experiment

In the main experiment, subjects had to indicate the modality of the stimulus, word or picture, by button-press. Subjects responded correctly in 97.5% of trials.

For replications of a same picture, cosine similarity between response patterns in the JA lesion VOI was significantly higher than chance (ACS = 0.074,  $P = 0.024$ ) (Fig. 4A; Table 2). For picture pairs belonging to a same semantic cluster, the cosine similarities between fMRI response patterns in the JA lesion VOI were also significantly higher than chance (ACS = 0.060,  $P < 0.005$ ) (Fig. 6A; Table 2). ACS between response patterns to pictures belonging to the musical instruments (ACS = 0.067) or vehicles (ACS = 0.069) clusters was higher than ACS between response patterns to tool pictures in the right fusiform VOI (ACS = 0.045). The effects of picture identity and of semantic cluster were also present in primary visual cortex (ACS = 0.057,  $P < 0.001$ ; ACS = 0.014,  $P = 0.001$ , respectively) (but see the invariance experiments for further characterization of this apparent semantic similarity effect) (Table 2).

With regards to the words, for replications of a same word (Fig. 4B) or for word pairs belonging to a same semantic cluster (Fig 6B), cosine similarity between fMRI response patterns was at chance level ( $P > 0.100$ ) (Table 2). For pairs of entities belonging to two different clusters, cosine similarity was far below significance for both pictures and words ( $P > 0.333$ ) (Table 2) (Fig. 6C,D).



**Figure 4** – Probability distributions for the effect of replicating a same entity. The red arrow indicates the average cosine similarity of the right fusiform activity patterns between replications of a same entity. **A.** Main experiment: replication of a same picture (total number of replications per picture = 2). **B.** Main experiment: replication of a same word (total number of replications per word = 2). **C.** First invariance experiment: replication of a same entity represented by a same picture with variations of location and orientation (total number of replications per entity = 8). **D.** Second invariance experiment: replication of a same entity represented by a same picture with variations of location, orientation, color and size (total number of replications per entity = 16). **E.** Exemplars experiment: replication of a same entity represented by pictures of different exemplars (total number of replications per entity = 8). X axis: cosine similarity averaged over the group of subjects. Y axis: absolute frequency of a given cosine similarity value across a total of 10,000 random permutation labellings. Dotted line: 95th percentile of the distribution.

	JA lesion VOI		Primary visual VOI	
	ACS	<i>P</i>	ACS	<i>P</i>
<b>Main experiment: pictures</b>				
Replication	<b>0.074</b>	0.024	<b>0.057</b>	<0.001
Within cluster	<b>0.060</b>	0.003	<b>0.014</b>	0.001
<b>First invariance experiment</b>				
Same entity	<b>0.011</b>	0.044	0.005	0.153
Within cluster	0.002	0.544	-0.001	0.622
Within location	<b>0.006</b>	0.027	<b>0.003</b>	0.030
<b>Second invariance experiment</b>				
Same entity	<b>0.010</b>	0.016	<b>0.006</b>	0.048
Within cluster	0.005	0.225	0.002	0.787
Within location	<b>0.007</b>	0.009	<b>0.010</b>	<0.001
<b>Exemplars experiment</b>				
Same entity	<b>0.008</b>	0.048	0.005	0.136
Within cluster	0.003	0.268	-0.002	0.591

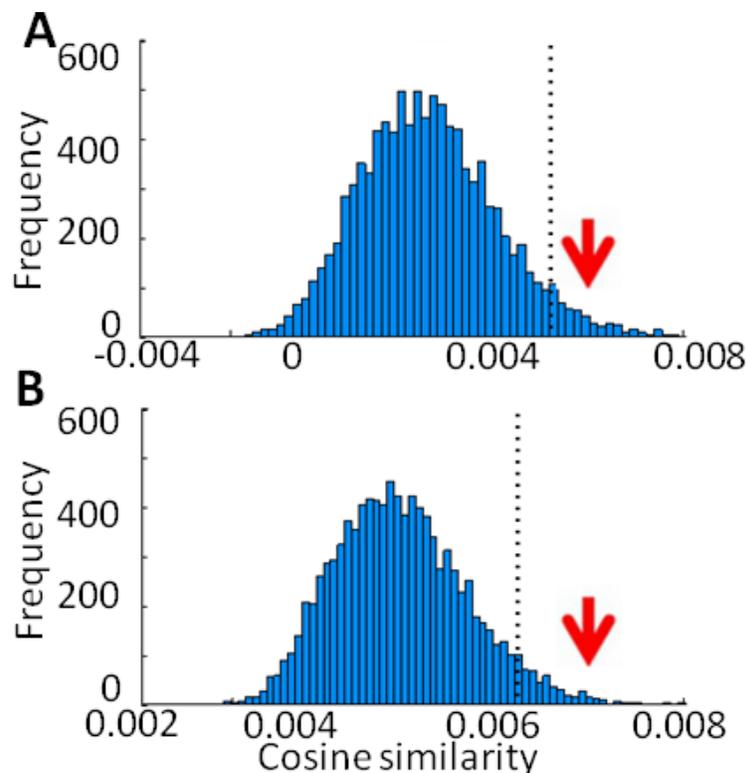
**Table 2** – Average Cosine Similarities (ACS) within the JA lesion VOI and the primary visual control VOI. Bold and underlined:  $P < 0.05$  Bonferroni corrected for number of comparisons; bold:  $P < 0.05$  uncorrected)

### 3.2. Invariance experiments

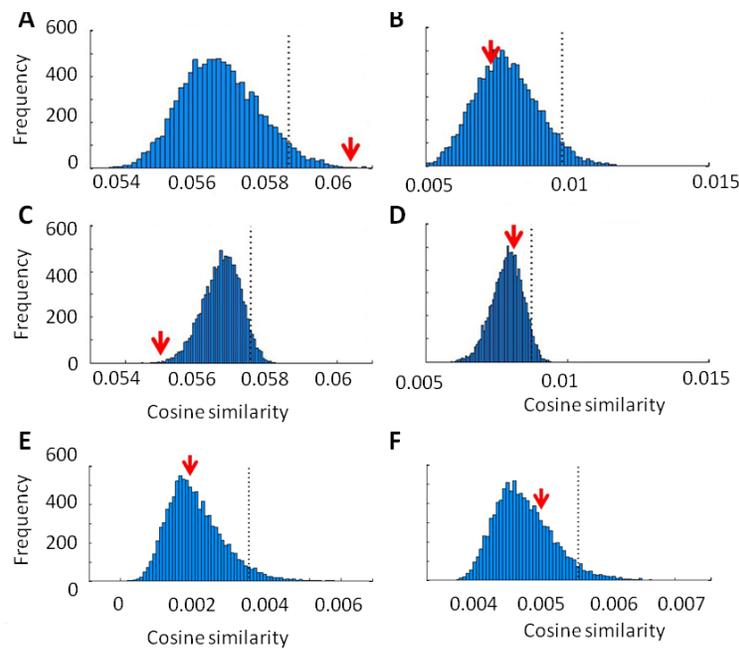
In the first invariance experiment, subjects had to fixate centrally and we varied the location of the stimuli. In the debriefing of the first invariance experiment, subjects were able to name 94.8% of stimuli when presented peripherally. Closely related answers (e.g. "violin" for "cello") were regarded as correct.

For replications of a same picture, cosine similarity between response patterns in the JA lesion VOI was significantly higher than chance despite variations in location and orientation (ACS = 0.011,  $P = 0.044$ ) (Fig. 4C; Table 2). The cosine similarity however within a semantic cluster was far below significance (ACS = 0.002,  $P = 0.544$ ) (Fig. 6E; Table 2). No effects of replication of a same picture or of semantic cluster were found in the primary visual control region when location was varied ( $P > 0.100$ ) (Table 2). In both the JA lesion VOI (Fig. 5A) and the primary visual control VOI, cosine similarities between same-location trials were significantly higher than chance (ACS = 0.006,  $P = 0.027$  and ACS = 0.003,  $P = 0.030$ , respectively) (Table 2).

In the second invariance experiment, the effects were essentially the same as in the first invariance experiment (e.g. replication of a same entity in JA lesion VOI: ACS=0.010,  $P = 0.016$  (Fig. 4D; same-location trials: ACS = 0.007,  $P = 0.009$ ) (Fig. 5B; Table 2; within-cluster ACS = 0.005,  $P > 0.2$  (Fig. 6F)).



**Figure 5** – First (A) and second (B) invariance experiment: Probability distributions for the effect of location. The red arrow indicates the average cosine similarity (ACS) of the right fusiform activity patterns for pictures presented in a same quadrant. X axis: cosine similarity averaged over the group. Y axis: absolute frequency of a given cosine similarity value across a total of 10,000 random permutation labellings. Dotted line: 95th percentile of the distribution.



**Figure 6** – Probability distributions for the effect of semantic cluster. The red arrow indicates the average cosine similarity (ACS) of the right fusiform activity patterns between entities belonging to a same (**A,B,E,F**) or a different (**C,D**) semantic cluster. **A**. Main experiment. ACS between picture pairs within semantic clusters. **B**. Main experiment. ACS between word pairs within semantic clusters. **C**. Main experiment. ACS between pairs of pictures belonging to two different semantic clusters. **D**. Main experiment. ACS between pairs of words belonging to two different semantic clusters. **E**. First invariance experiment. ACS between pairs of pictures belonging to a same semantic cluster with variations in parafoveal location and orientation. **F**. Second invariance experiment. ACS between pairs of pictures belonging to a same semantic cluster with variations in location, size, color and orientation. X axis: cosine similarity averaged over the group. Y axis: absolute frequency of a given cosine similarity value across a total of 10,000 random permutation labellings. Dotted line: 95th percentile of the distribution.

### 3.3. Exemplars experiment

In the exemplars experiment subjects were able to name 99.5% of stimuli correctly. When a same entity was represented as different exemplars, cosine similarities between responses to exemplars of a same entity were significantly higher than chance in the JA lesion volume (ACS = 0.008,  $P = 0.048$ ) (Fig. 4E), without effects of semantic cluster (Table 2). No such effects were seen in the primary visual VOI (ACS = 0.005,  $P > 0.100$ ).

### 3.4. Ancillary analyses

To evaluate how specific our findings were for the right as opposed to the left posterior/middle fusiform cortex, we mirrored the JA volume to the left and conducted identical analyses of the homotopical left-sided activity patterns. Overall, the pattern in the left mirror volume was comparable to that seen to the right. In the main experiment, for replications of a same picture, ACS between response patterns was significantly above chance in the left fusiform mirror volume (ACS = 0.0610,  $P = 0.030$ ). For pictures belonging to a same semantic cluster, ACS was also significantly above chance (ACS = 0.050,  $P = 0.001$ ). No effects were found for words (replication: ACS = 0.017,  $P = 0.188$ ; semantic cluster: ACS = 0.010,  $P = 0.411$ ). In the first invariance experiment, in the left-sided mirror region, both the effect of object identity and the effect of semantic similarity disappeared when location was varied. For entities presented in the same quadrant of the visual field, ACS was increased (ACS = 0.005,  $P = 0.009$ ). In the second invariance experiment, in the left-sided mirror volume cosine similarity was increased for replications of a same entity despite variations in location, color, size and orientation (ACS = 0.009,  $P = 0.006$ ), without semantic similarity effects. This was also the case in the exemplars experiment (ACS = 0.011,  $P = 0.008$ ).

## 4. Discussion

The two key findings in the right fusiform lesion-based volume (Fig. 1) are, first, that similarity between activity patterns reflects object identity in a manner that is invariant for location, size, and color (color versus grayscale) and also generalizes across exemplars (Table 2, Fig. 4). Second, similarity between activity patterns was originally seen within semantic clusters but this proved spurious when invariance for location was added as a criterion (Table 2, Fig. 6). Our results fit with our a priori hypothesis that invariant processing of object identity is dissociable from the processing of the conceptual similarity between objects and that right posterior/middle fusiform cortex plays a critical role in the former process.

Our main research question was based on the behavioral deficit induced by the structural lesion in JA (Vandenbulcke et al., 2006). Of course, within a relatively large VOI there may be inter-individual variability in the pattern coding for object identity or semantic similarity. Since we compare the similarity of the entire pattern between different stimuli within each subject, this type of between-subject variability is accounted for. We used cosine similarity, but other similarity measures (e.g. Jaccard similarity coefficient, Pearson correlation) would have been equally valid (Mitchell et al., 2008). One of the remarkable strengths of our approach is that it can estimate the similarity between as few as two picture replications among a set of 60 different objects (Fig. 4A). The magnitude of the cosine similarity values as such may appear to be relatively low. When comparing between studies, it is important to ensure that similarity values are based on the same mathematical formulation. For instance, Weber et al. (2009) report higher Pearson correlation values than we obtained, but these pertain to a correlation between fMRI and behavioral data (second-order similarity). To our knowledge, only Mitchell et al. (2008) has reported cosine similarity values in a comparable manner. In their study, the values are of a similar order of magnitude (e.g. .17 as compared to .07 for picture replication), be it somewhat higher Mitchell et al. (2008). This difference can be accounted for by the fact that we did not select voxels based on stability in contrast to Mitchell et al. (2008) (i.e. feature selection). Stability refers to a high correlation of activity patterns between multiple trial presentations and if used as a selection criterion, this is bound to increase the absolute values for the cosine similarity between these trials.

In the main experiment, we initially found increased cosine similarity between response patterns to pictures belonging to the same semantic cluster (Fig. 6A). The apparent semantic similarity effect disappeared once we varied the location, size-on-the-screen or color versus grayscale (Table 2; Fig. 6E,F). Our semantic similarity findings are in line with previous fMRI studies in healthy volunteers based on univariate analyses that have attributed differences between semantic categories in this region to latent perceptual confounds (Price and Devlin, 2003; Devlin et al., 2005; Bright et al., 2005). Caution needs to be exercised when MVPA is applied to study category effects as subtle perceptual differences between classes of objects may be picked up by this highly sensitive method. Visuo-perceptual similarity between entities from a same category is often intertwined with category membership, at least if stimuli are only presented through the visual input-modality and in particularly so when broad categories such as animate versus inanimate entities are compared (Devlin et al., 2002; Forde et al., 1997; Gale et al., 2001; Garrard et al., 2001; Mahon and Caramazza, 2009; Op de Beeck et al., 2008; Pilgrim et al., 2005; Pulvermüller et al., 2009; Rogers and McClelland, 2004; Sartori and Job, 1988). Some previous MVPA studies of semantic category effects (e.g. Cox and Savoy (2003); Eger et al. (2008); Haxby et al. (2001); Kriegeskorte et al. (2008)) did not manipulate location. Most MVPA studies (e.g. (Cox and Savoy, 2003; Eger et al., 2008; Haxby et al., 2001; Kriegeskorte et al., 2008; Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008) contrast broad categories such as animate versus inanimate entities. Even if the experimenter deliberately attempts to match stimuli from such widely different categories for perceptual differences, subtle differences that are not strictly semantic in nature may remain, e.g. category-dependent differences in perceptual similarity between items (Price and Devlin, 2003; Devlin et al., 2005; Bright et al.,

2005). For that reason, we selected entities that were semantically close, but distinctive enough to be assigned to separate semantic clusters based on behavioral measures (Fig. 3A). Musical instruments and vehicles were matched for perceptual and lexical variables. Tools however were visually less complex (Table 1). Our MVPA semantic similarity effects remained essentially the same if the tools cluster was removed. The lack of invariance however of the semantic similarity effect strongly suggests that the apparent semantic similarity effect was due to hidden perceptual similarities between the pictures, despite the prior matching and despite the fact that the categories studied were closely related.

The current findings amount to a single dissociation: basic-level object identity is reflected in the right fusiform response patterns (Fig. 4) but we did not find an invariant effect of semantic similarity (Fig. 6E,F). It will be important to evaluate whether semantic similarity effects can be obtained using the same procedures e.g. in more anterior regions (Sharp et al., 2004; Mion et al., 2010; Peelen and Caramazza, 2012). Surface electrode studies have revealed category selectivity in anterior temporal cortex (Liu et al., 2009; Chan et al., 2011). In a recent study, Peelen and Caramazza (2012) evaluated the effect of category membership (based on action (squeezing versus rotating) or location (kitchen versus garage)) on activity patterns and reported effects mainly in anterior temporal cortex. In the same study, Peelen and Caramazza (2012) constructed a perceptual similarity matrix between the pictures based on pairwise judgment of perceptual similarity performed offline by the participants. The perceptual similarity effects were mainly found in posterior fusiform cortex (Peelen and Caramazza, 2012), an area which did not show category effects. While the basic research question is closely similar to that addressed by Peelen and Caramazza (2012), our experimental approach differs in several aspects. Most notably, we used a larger set of stimuli and more semantic clusters (Fig. 3), tested predictions based on JA's neuropsychological deficit (Fig. 1), and checked invariance as a criterion for object identity and semantic similarity effects (Fig. 4C-E, 6E-F). Furthermore, subjects had to perform a shallow task or viewed the stimuli passively whereas they performed a one-back semantic decision task in the Peelen and Caramazza (2012) study. Despite these differences, our findings in posterior and mid-fusiform gyrus are in accordance with Peelen and Caramazza (2012).

Our MVPA findings are in line with the invariance of object processing in right fusiform cortex for location (Grill-Spector et al., 1999), size (Grill-Spector et al., 1999) and viewpoint (Konen and Kastner, 2008) as demonstrated by fMRI-adaptation. Invariance is a key characteristic of shape processing in LO (Kanwisher, 1997; Liu et al., 2008; Malach et al., 1995; Schacter et al., 1995). In the monkey brain, relative shape preference in IT is invariant for cue (luminance, relative motion or texture), size or location (Vogels and Orban, 1996). Position-tolerant identity information can be derived from inferior temporal (IT) neuronal activity (Hung et al., 2005; Tanaka, 2003). According to our MVPA analysis, location as well as location-invariant object identity are coded within a same region, in accordance with Cichy et al. (2011) and Schwarzlose et al. (2008). This dual coding fits with the characteristics of inferotemporal neurons as demonstrated in non-human primates: Inferotemporal neurons show selectivity for object location (Op De Beeck and Vogels, 2000) and are also selective for object identity in a location-invariant manner (Op De Beeck and Vogels, 2000; Rust and Dicarlo, 2010; DiCarlo and Cox, 2007; Liu et al., 2009; DiCarlo et al., 2012).

To conclude, identity of individual objects can be derived from right mid-posterior fusiform activity patterns even for large sets of closely related objects with varying location, size and color. No proper semantic similarity effect, however, could be found in this region once invariance was added as a criterion. Our MVPA findings place the contribution of right mid-posterior fusiform cortex at the level of processing structural descriptions (Forde et al., 1997; Humphreys et al., 1997; Humphreys and Riddoch, 2003), in line with previous lesion studies (Vandenbulte et al., 2006; Konen et al., 2011). Further studies will be needed to localize invariant effects of semantic similarity, with as a prime candidate anteromedial temporal cortex (Liu et al., 2009; Mion et al., 2010; Chan et al., 2011; Peelen and Caramazza, 2012).

## 5. References

- Baayen, R.H., Piepenbrock, R., van Rijn, H., 1993. The CELEX lexical database. (CD-ROM) Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Barsalou, L.W., 2008. Grounded cognition. *Annu Rev Psychol* 59, 617–645.
- Behrmann, M., Kimchi, R., 2003a. Perceptual Organization in Vision: Behavioral and neural perspectives. Hove, UK: Lawrence-Erlbaum Associates. chapter Visual Perceptual Organization: Lessons from Lesions. pp. 337–376.
- Behrmann, M., Kimchi, R., 2003b. What does visual agnosia tell us about perceptual organization and its relationship to object perception? *J Exp Psychol Hum Percept Perform* 29, 19–42.
- Behrmann, M., Peterson, M.A., Moscovitch, M., Suzuki, S., 2006. Independent representation of parts and the relations between them: evidence from integrative agnosia. *J Exp Psychol Hum Percept Perform* 32, 1169–1184.
- Behrmann, M., Williams, P., 2007. Impairments in part-whole representations of objects in two cases of integrative visual agnosia. *Cogn Neuropsychol* 24, 701–730.
- Biederman, I., 1987. Recognition-by-components: a theory of human image understanding. *Psychol Rev* 94, 115–147.
- Bright, P., Moss, H.E., Stamatakis, E.A., Tyler, L.K., 2005. The anatomy of object processing: the role of anteromedial temporal cortex. *Q J Exp Psychol B* 58, 361–377.
- Chan, A.M., Baker, J.M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., Cash, S.S., Halgren, E., 2011. First-pass selectivity for semantic categories in human anteroventral temporal lobe. *J Neurosci* 31, 18119–18129.
- Cichy, R.M., Chen, Y., Haynes, J.D., 2011. Encoding the identity and location of objects in human LOC. *Neuroimage* 54, 2297–2307.
- Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* 19, 261–270.
- De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M., Voorspoels, W., Storms, G., 2008. Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behav Res* 40, 1030–1048.
- Dell, G.S., 1986. A spreading-activation theory of retrieval in sentence production. *Psychol Rev* 93, 283–321.
- Devlin, J.T., Rushworth, M.F.S., Matthews, P.M., 2005. Category-related activation for written words in the posterior fusiform is task specific. *Neuropsychologia* 43, 69–74.
- Devlin, J.T., Russell, R.P., Davis, M.H., Price, C.J., Moss, H.E., Fadili, M.J., Tyler, L.K., 2002. Is there an anatomical basis for category-specificity? Semantic memory studies in PET and fMRI. *Neuropsychologia* 40, 54–75.
- DiCarlo, J.J., Cox, D.D., 2007. Untangling invariant object recognition. *Trends Cogn Sci* 11, 333–341.

- DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object recognition? *Neuron* 73, 415–434.
- Eger, E., Ashburner, J., Haynes, J.D., Dolan, R.J., Rees, G., 2008. fMRI activity patterns in human LOC carry information about object exemplars within category. *J Cogn Neurosci* 20, 356–370.
- Farah, M., 2004. *Visual Agnosia*. Cambridge, MA, Massachusetts Institute of Technology Press.
- Forde, E., Francis, D., Riddoch, M., Rumiati, R., Humphreys, G., 1997. On the links between visual knowledge and naming: a single case study of a patient with a category-specific impairment for living things. *Cogn Neuropsychol* 14(3), 403–458.
- Gale, T.M., Done, D.J., Frank, R.J., 2001. Visual crowding and category specific deficits for pictorial stimuli: A neural network model. *Cogn Neuropsychol* 18, 509–550.
- Garrard, P., Ralph, M.A., Hodges, J.R., Patterson, K., 2001. Prototypicality, distinctiveness, and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Cogn Neuropsychol* 18, 125–174.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., Malach, R., 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24, 187–203.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- Hillis, A., Caramazza, A., 1995. Cognitive and Neural Mechanisms Underlying Visual and Semantic Processing: Implications from Optic Aphasia. *Journal of Cognitive Neuroscience* 4:7, 457–478.
- Humphreys, G., Riddoch, J., 2003. *The organisation of conceptual knowledge in the brain*. Hove, UK: Psychology Press.
- Humphreys, G.W., Riddoch, M.J., Price, C.J., 1997. Top-down processes in object identification: evidence from experimental psychology, neuropsychology and functional anatomy. *Philos Trans R Soc Lond B Biol Sci* 352, 1275–1282.
- Hung, C.P., Kreiman, G., Poggio, T., DiCarlo, J.J., 2005. Fast readout of object identity from macaque inferior temporal cortex. *Science* 310, 863–866.
- Huth, A., Nishimoto, S., Vu, A., Gallant, J., 2012. A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron* 76, 1210–1224.
- Kanwisher, N., 1997. A locus in human extrastriate cortex for visual shape analysis. *J Cogn Neurosci* 9:1, 133–142.
- Kayaert, G., Biederman, I., Vogels, R., 2003. Shape tuning in macaque inferior temporal cortex. *J Neurosci* 23, 3016–3027co.
- Konen, C.S., Behrmann, M., Nishimura, M., Kastner, S., 2011. The functional neuroanatomy of object agnosia: a case study. *Neuron* 71, 49–60.
- Konen, C.S., Kastner, S., 2008. Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci* 11, 224–231.

- Kourtzi, Z., Kanwisher, N., 2001. Representation of perceived object shape by the human lateral occipital complex. *Science* 293, 1506–1509.
- Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci* 2, 4.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., Malach, R., 2001. A hierarchical axis of object processing stages in the human visual cortex. *Cereb Cortex* 11, 287–297.
- Liu, H., Agam, Y., Madsen, J.R., Kreiman, G., 2009. Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62, 281–290.
- Liu, X., Steinmetz, N.A., Farley, A.B., Smith, C.D., Joseph, J.E., 2008. Mid-fusiform activation during object discrimination reflects the process of differentiating structural descriptions. *J Cogn Neurosci* 20, 1711–1726.
- Logothetis, N.K., Sheinberg, D.L., 1996. Visual object recognition. *Annu Rev Neurosci* 19, 577–621.
- Mahon, B.Z., Caramazza, A., 2009. Concepts and categories: a cognitive neuropsychological perspective. *Annu Rev Psychol* 60, 27–51.
- Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci U S A* 92, 8135–8139.
- Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19, 1233–1239.
- Marr, D., 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco (CA), USA: W. H. Freeman.
- Martin, A., Chao, L.L., 2001. Semantic memory and the brain: structure and processes. *Curr Opin Neurobiol* 11, 194–201.
- Martin, A., Ungerleider, L., Haxby, J., 1999. *The New Cognitive Neurosciences: Second Edition*. Cambridge, MA, Massachusetts Institute of Technology Press. chapter Category specificity and the brain: the sensory/motor model of semantic representations of objects. pp. 1023–1036.
- Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y.T., Fryer, T.D., Williams, G.B., Hodges, J.R., Nestor, P.J., 2010. What the left and right anterior fusiform gyri tell us about semantic memory. *Brain* 133, 3256–3268.
- Mitchell, T.M., Shinkareva, S.V., Carlson, A., Chang, K.M., Malave, V.L., Mason, R.A., Just, M.A., 2008. Predicting human brain activity associated with the meanings of nouns. *Science* 320, 1191–1195.
- Nobre, A.C., Allison, T., McCarthy, G., 1994. Word recognition in the human inferior temporal lobe. *Nature* 372, 260–263.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Op de Beeck, H., 2010. Probing the mysterious underpinnings of multi-voxel fMRI analyses. *Neuroimage* 50, 567–571.

- Op de Beeck, H., Haushofer, J., Kanwisher, N., 2008. Interpreting fMRI data: maps, modules and dimensions. *Nat Rev Neurosci* 9, 123–135.
- Op De Beeck, H., Vogels, R., 2000. Spatial sensitivity of macaque inferior temporal neurons. *J Comp Neurol* 426, 505–518.
- Peelen, M.V., Caramazza, A., 2012. Conceptual object representations in human anterior temporal cortex. *J Neurosci* 32, 15728–15736.
- Pilgrim, L.K., Moss, H.E., Tyler, L.K., 2005. Semantic processing of living and nonliving concepts across the cerebral hemispheres. *Brain Lang* 94, 86–93.
- Price, C.J., Devlin, J.T., 2003. The myth of the visual word form area. *Neuroimage* 19, 473–481.
- Pulvermüller, F., Kherif, F., Hauk, O., Mohr, B., Nimmo-Smith, I., 2009. Distributed cell assemblies for general lexical and category-specific semantic processing as revealed by fMRI cluster analysis. *Hum Brain Mapp* 30, 3837–3850.
- Riddoch, M., Humphreys, G., 1993. *Birmingham Object Recognition Battery*. Hove, UK: Lawrence Erlbaum Associates Ltd.
- Rogers, T., McClelland, J., 2004. *Semantic cognition. A parallel distributed approach*. Cambridge, MA, Massachusetts Institute of Technology Press.
- Rumiati, R., Humphreys, G., 1997. Visual Object Agnosia without Alexia or Prosopagnosia: Arguments for Separate Knowledge Stores. *Visual Cognition* 4:2, 207217.
- Rust, N.C., Dicarlo, J.J., 2010. Selectivity and tolerance (“invariance”) both increase as visual information propagates from cortical area V4 to IT. *J Neurosci* 30, 12978–12995.
- Sartori, G., Job, R., 1988. The oyster with four legs: a neuropsychological study on the interaction between vision and semantic information. *Cogn. Neuropsychol.* 5, 105–132.
- Sayres, R., Grill-Spector, K., 2008. Relating retinotopic and object-selective responses in human lateral occipital cortex. *J Neurophysiol* 100, 249–267.
- Schacter, D.L., Reiman, E., Uecker, A., Polster, M.R., Yun, L.S., Cooper, L.A., 1995. Brain regions associated with retrieval of structurally coherent visual information. *Nature* 376, 587–590.
- Schwarzlose, R.F., Swisher, J.D., Dang, S., Kanwisher, N., 2008. The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci U S A* 105, 4447–4452.
- Sharp, D.J., Scott, S.K., Wise, R.J.S., 2004. Retrieving meaning after temporal lobe infarction: the role of the basal language area. *Ann Neurol* 56, 836–846.
- Tanaka, K., 2003. Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cereb Cortex* 13, 90–99.
- Tootell, R.B., Hadjikhani, N.K., Mendola, J.D., Marrett, S., Dale, A.M., 1998. From retinotopy to recognition: fMRI in human visual cortex. *Trends Cogn Sci* 2, 174–183.
- Tyler, L., Moss, H., 2001. Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences* 5, 244–252.

- Van Essen, D., 2005. A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage* 28, 635–662.
- Vandenbulcke, M., Peeters, R., Fannes, K., Vandenberghe, R., 2006. Knowledge of visual attributes in the right hemisphere. *Nat Neurosci* 9, 964–970.
- Vogels, R., Orban, G.A., 1996. Coding of stimulus invariances by inferior temporal neurons. *Prog Brain Res* 112, 195–211.
- Wager, T.D., Nichols, T.E., 2003. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. *Neuroimage* 18, 293–309.
- Weber, M., Thompson-Schill, S.L., Osherson, D., Haxby, J., Parsons, L., 2009. Predicting judged similarity of natural categories from their neural representations. *Neuropsychologia* 47, 859–868.