

**Left perirhinal cortex codes for similarity in meaning between written words:
Comparison with auditory word input**

Antonietta Gabriella Liuzzi^a, Rose Bruffaerts^a, Patrick Dupont^a, Katarzyna Adamczuk^a,
Ronald Peeters^c, Simon De Deyne^d, Gerrit Storms^d, Rik Vandenberghe^{a,b}

^aLaboratory for Cognitive Neurology, Department of Neurosciences, University of Leuven, Belgium;
^bNeurology Department, University Hospitals Leuven, Belgium; ^cRadiology Department, University
Hospitals Leuven, Leuven, Belgium. ^dLaboratory of Experimental Psychology, Humanities and Social
Sciences Group, University of Leuven, Leuven, Belgium.

Correspondence to: Rik Vandenberghe. Neurology Department, University Hospitals Leuven, Her-
estraat 49 bus 7003, B-3000 Leuven, Belgium. Tel: +3216344280, Fax: +3216344285. E-mail:
`rik.vandenberghe@uz.kuleuven.ac.be`.

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in Neuropsychologia following peer review. The definitive publisher-authenticated version (Liuzzi A.G., Bruffaerts R., Dupont P., Adamczuk K., Peeters R., De Deyne S., Storms G., Vandenberghe R.(2015) Left perirhinal cortex codes for similarity in meaning between written words: Comparison with auditory word input.) is available online at: <http://dx.doi.org/10.1016/j.neuropsychologia.2015.03.016>

Abstract

Left perirhinal cortex has been previously implicated in associative coding. According to a recent experiment, the similarity of perirhinal fMRI response patterns to written concrete words is higher for words which are more similar in their meaning. If left perirhinal cortex functions as an amodal semantic hub, one would predict that this semantic similarity effect would extend to the spoken modality. We conducted an event-related fMRI experiment and evaluated whether a same semantic similarity effect could be obtained for spoken as for written words. Twenty healthy subjects performed a property verification task in either the written or the spoken modality. Words corresponded to concrete animate entities for which extensive feature generation was available from more than 1000 subjects. From these feature generation data, a concept-feature matrix was derived which formed the basis of a cosine similarity matrix between the entities reflecting their similarity in meaning (called the "semantic cossimilarity matrix"). Independently, we calculated a cosine similarity matrix between the left perirhinal fMRI activity patterns evoked by the words (called the "fMRI cossimilarity matrix"). Next, the similarity was determined between the semantic cossimilarity matrix and the fMRI cossimilarity matrix. This was done for written and spoken words pooled, for written words only, for spoken words only, as well as for crossmodal pairs. Only for written words did the fMRI cossimilarity matrix correlate with the semantic cossimilarity matrix. Contrary to our prediction, we did not find any such effect for auditory word input nor did we find cross-modal effects in perirhinal cortex between written and auditory words. Our findings situate the contribution of left perirhinal cortex to word processing at the top of the visual processing pathway, rather than at an amodal stage where visual and auditory word processing pathways have already converged.

Keywords: anterior temporal, fMRI, semantic, reading, MVPA, orthographic, spoken

1 Introduction

Although semantic processing from different input modalities is a fundamental issue in cognitive neuroscience, there is no agreement yet about the precise contribution of the different brain regions belonging to the associative-semantic network, such as ventral occipitotemporal, posterior or anterior temporal or inferior frontal cortex, let alone the differential contribution of sub-regions within these association zones. Also, little is known about how concrete entities are represented into such semantic network and how this is influenced by the input-modality of entities it processes.

For many years, neuroimaging studies of the processing of meaning of words and pictures have relied on univariate analyses (for review see Binder et al. (2009)). These studies suggested possible candidate regions for amodal processing of words and pictures, such as the ventral occipitotemporal cortex (Chao et al., 1999; Buckner et al., 2000; Seghier and Price, 2011), the posterior middle temporal gyrus (Chertkow et al., 1997; Vandenberghe et al., 2007; Kircher et al., 2009), the angular gyrus (Bonner et al., 2013), the anterior temporal pole (Hodges et al., 1992; Vandenberghe et al., 1996; Rogers and McClelland, 2004; Patterson et al., 2007; Binder et al., 2009; Lambon Ralph et al., 2010), the left inferior frontal sulcus (Wagner et al., 1997), the anterior inferior frontal gyrus (Goldberg et al., 2007), and the left middle frontal gyrus (Demb et al., 1995; Vandenberghe et al., 1996; Van Doren et al., 2010). Part of the amodality, especially in left inferior frontal sulcus and ventral occipitotemporal transition zone, may be due to working memory (Badre et al., 2005; Van Doren et al., 2010) or executive control (Whitney et al., 2011) which is required by task regardless of input-modality. Other regions may be involved in the representation of concrete entities regardless of the input-modality.

The widespread application of multivariate pattern analysis (MVPA) techniques has heralded a new research era in the study of the representation of concrete entities in the human brain. An ever increasing number of studies have succeeded in detecting patterns of brain activity that reflect the identity or the semantic category of pictures (Haxby et al., 2001; Hanson et al., 2004; Shinkareva et al., 2008; Connolly et al., 2012; Mur et al., 2012; Bruffaerts et al., 2013a), written words (Mitchell et al., 2008; Just et al., 2010), written words and pictures (Shinkareva et al., 2011; Simanova et al., 2012; Bruffaerts et al., 2013b; Devereux et al., 2013; Fairhall and Caramazza, 2013), written and spoken words (Akama et al., 2012; Simanova et al., 2012) as well as natural sounds (Simanova et al., 2012).

Left perirhinal cortex is one of the regions that has come out of a number of MVPA studies of this kind in a relatively consistent manner (Bruffaerts et al., 2013b; Fairhall and Caramazza, 2013; Clarke and Tyler, 2014; Mur, 2014). For written words, semantic similarity between concrete animate entities is reflected by the cosine similarity between fMRI response patterns in perirhinal cortex (Bruffaerts et al., 2013b). The semantic similarity effect in perirhinal cortex was subsequently confirmed by Clarke and Tyler (2014) during overt naming of pictures of animate and inanimate items. A third research group independently found crossmodal semantic effects for written words and pictures in left ventral temporal cortex, encompassing left perirhinal cortex among other areas (Fairhall and Caramazza, 2013).

The primary objectives of the current study were to evaluate the replicability of our left

perirhinal effect for written words in an independent dataset, and secondly to determine whether the effect was also present for spoken modality. The latter question is important: if the semantic similarity effect extends to spoken words, this would suggest that left perirhinal cortex may be one of the amodal hubs (Visser et al., 2010b; Visser and Lambon Ralph, 2011) processing concrete entities at a more abstract level regardless of the input-modality of the referent.

2 Subjects and Methods

2.1 Participants

Twenty subjects, between 18 and 28 years, participated in this fMRI experiment. All subjects were native Dutch speakers, right-handed, free of neurological history, psychiatric history and had no hearing impairment. All the procedures were approved by the Ethics Committee of the University Hospital of Leuven.

2.2 Stimuli and task

2.2.1 Stimuli and semantic cosine similarity matrix

The stimulus set and the properties used for property verification were identical to those used by Bruffaerts et al. (2013b). Stimuli and properties were derived from a dataset including a large number of animate concepts collected by De Deyne et al. (2008). In the De Deyne et al. (2008) study, 1003 students were asked to write down 10 features for each of the 10 concrete nouns they received, each noun corresponding to a concept (a concrete entity) (feature generation task). A minimum of 180 features for each concept were collected, for a total of more than 120 concepts. Next, a feature applicability judgment task was performed by 4 subjects who were asked, for each concept-feature pair, to judge whether the feature was applicable to the concept (De Deyne et al., 2008). In this way a feature applicability matrix (or concept-feature matrix) was generated and all properties were verified for all entities. This matrix contains one column per entity (for a total of more than 120 entities) and one row per property generated by the more than 1000 subjects in response to these entities, yielding a total of 764 rows. Each cell of this matrix contains the sum of the positive responses by 4 subjects who judged the applicability of a given property for a given entity (integer value between 0 and 4) (De Deyne et al., 2008).

For use in the current experiment, 24 animate entities from 6 semantic subcategories (birds, marine animals, fish, herpetofauna, insects, and livestock) were selected based on hierarchical clustering analysis (Fig 1) (Bruffaerts et al., 2013b). Between the 6 semantic subcategories, lexical parameters (word frequency, (Baayen et al., 1993) familiarity; age of acquisition and image-ability) were matched as much as possible (Cree and McRae, 2003; De Deyne et al., 2008; Bruffaerts et al., 2013b). For a statistical comparison of these variables between the subcategories, see Table 1 in Bruffaerts et al. (2013b). Written word length was between 3 and 11 characters. The written words were presented with a letter size of 0.7 visual degrees. Spoken words intensity was fixed at 35.4 dB. The duration of the spoken words was between 0.33 and

1.5s (birds: mean = 0.95, SD = 0.102; marine animals: mean = 0.96, SD = 0.278; fish: mean = 1.03, SD = 0.319; herpetofauna: mean = 1.18, SD = 0.301; insects: mean = 0.94, SD = 0.14; livestock: mean = 0.62, SD = 0.24).

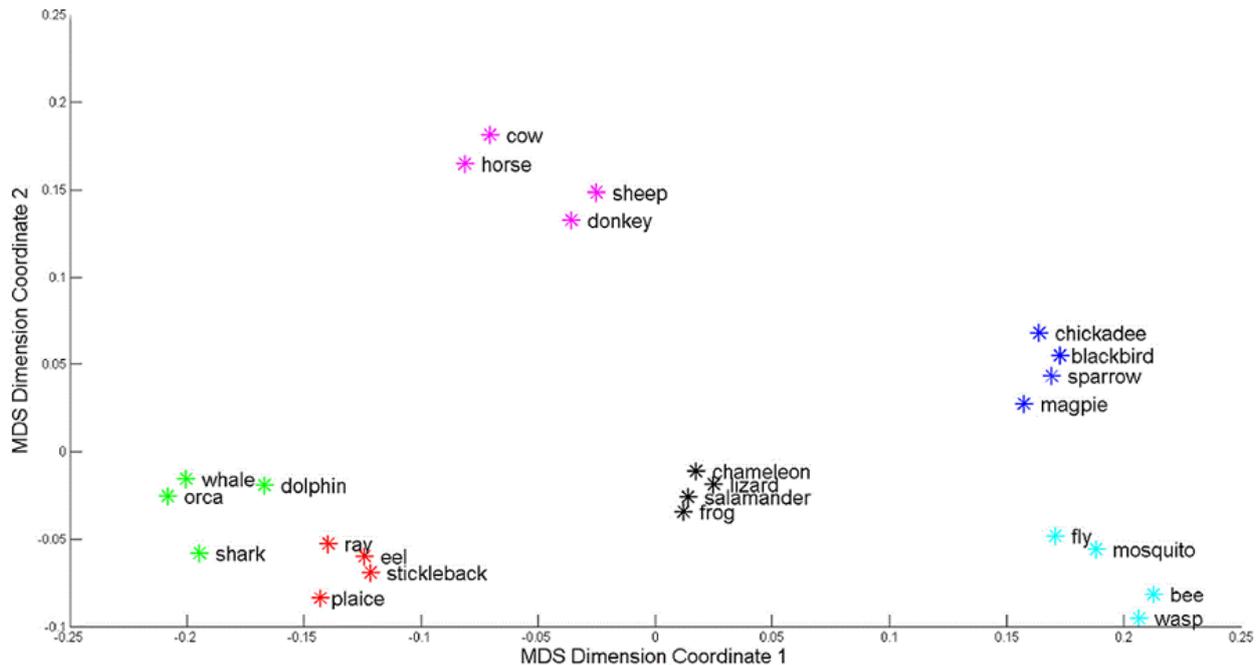


Fig. 1 Visual representation of the semantic clusters and semantic distances between entities, based on the feature generation data collected by Deyne and Storms (2008). For visualization, data reduction of the similarity matrix to two dimensions was performed by means of multidimensional scaling (MDS) (Matlab 2011b, Statistics Toolbox). The 6 semantic clusters correspond to livestock (magenta), birds (blue), insects (cyan), fish (red), herpetofauna (black) and large marine animals (green).

From the feature applicability matrix we selected the 24 columns corresponding to the 24 entities used in the current fMRI experiment. We used the entire column (i.e. 764 rows, one row per property). Then, we calculated the pairwise cosine similarity (CS) between each pair of columns (entities) and this yielded the semantic cosine similarity matrix for the entities used in our fMRI study.

Next, we evaluated whether other variables than semantic ones influenced the semantic cossimilarity values. We examined how orthographic distance related to the semantic cossimilarity matrix. We calculated the orthographic distance matrix for the 24 words based on the Levenshtein distances. The Levenshtein distances are commonly used as a measure to quantify orthographic distance in psycholinguistics (Yarkoni et al., 2008) and corresponds to the minimal number of steps (additions, removals, replacements) that one has to undertake to transform an orthographic string into the other string. We transformed the semantic cosine similarity matrix into a semantic cosine distance matrix by subtracting the semantic cosine similarity from 1. The cosine similarity between the orthographic distance matrix and the semantic cosine distance matrix was not significant ($P > 0.8$). In the same way, we evaluated whether two orthographic variables, word length and orthographic neighborhood size (based on orthographic Levenshtein distance to the nearest 20 neighbours (OLD20) (Yarkoni et al., 2008)), were related with the semantic cosine distance matrix. No effects were seen for either of the two orthographic variables ($P > 0.68$).

We also evaluated whether word frequency, age-of-acquisition (AoA), imageability, and word familiarity had an effect on the semantic cosine dissimilarity matrix. For each of these variables we calculated a 24-by-24 matrix containing per cell the pairwise difference in the value. For these 4 variables, the cosine similarity between the semantic cosine distance matrix and the difference matrix was significant, meaning that a larger pairwise difference corresponded to a higher semantic dissimilarity: word frequency (CS = 0.74, $P < 0.0006$), AoA (CS = 0.81, $P = 0.016$), imageability (CS = 0.79, $P = 0.036$) and word familiarity (CS = 0.83, $P = 0.0015$).

2.2.2 Task

In the fMRI, the participants performed a property verification task (Fig 2). During this task subjects were asked to judge whether or not a given animal possessed a given property. Subjects were instructed to keep the response box in their right hand. They had a response box in their right hand and pressed a lower or upper button depending on a "yes" or "no" judgment. This was counterbalanced between subjects. Subjects were told to respond as soon as they had made a decision and to guess if they were uncertain. For each property, the correct response was positive for approximately half of the entities. The eight properties involved in the task ("large", "legs", "wings", "smooth", "exotic", "mammal", "sea" and "flies") were selected from the most frequently generated by the 1003 students. The properties were presented as fully grammatical questions, e.g. "is it exotic?"

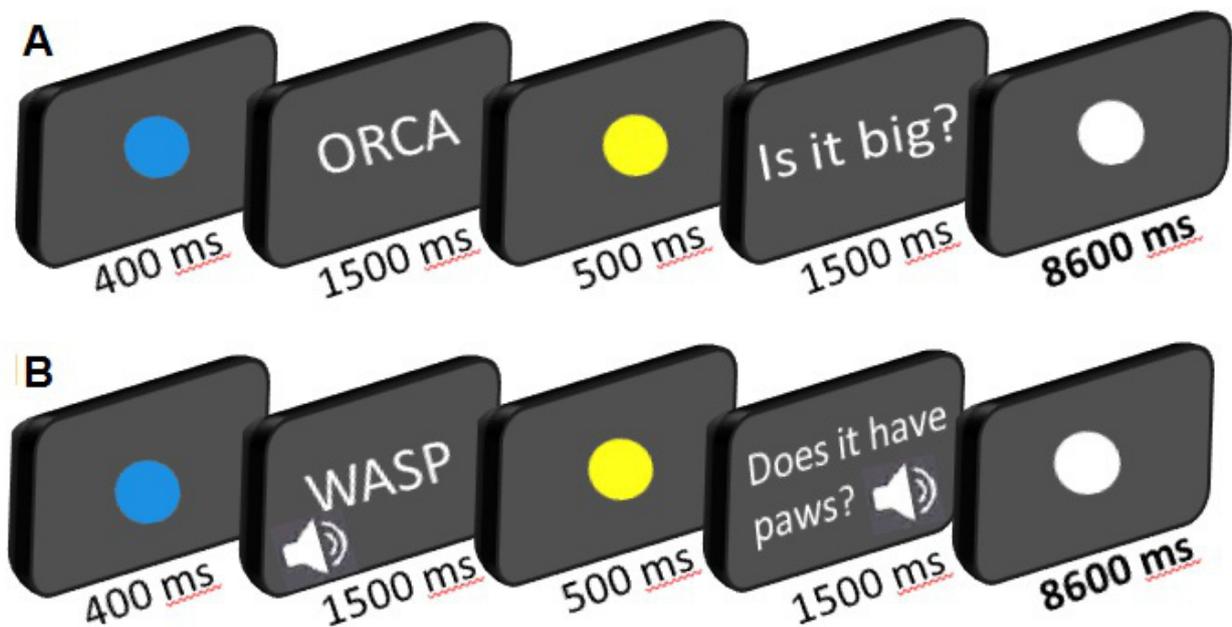


Fig. 2 Experimental task: (A) Property verification task in written modality. (B) Property verification task in spoken modality. In spoken modality both the word and question have been presented in spoken modality

2.2.3 Two trial types

The experiment contained two types of trials: the stimulus and the probe question were presented in either the written or the spoken modality. At the start of each trial the fixation point turned from white to blue. The blue fixation point remained on the screen for 400 ms, next

the fixation point disappeared and at the same time the sample stimulus was presented. In the visual condition, the written word was on the screen for 1500 ms. In the auditory condition the word was auditorily presented. A total of 1500 ms after the onset of the sample stimulus, a yellow fixation point appeared which remained on the screen for 500 ms. Next, the fixation point disappeared for 1500 ms and at the same time the property verification question was presented (Fig. 2). In the visual condition the property verification question was on the screen for 1500 ms. In the auditory condition the question was presented auditorily. A total of 1500 ms after the onset of the property verification question, the fixation point re-appeared. The interval from the re-appearance of the fixation point till the onset of the next trial was 8600 ms. Each trial therefore lasted in total for 12,500 ms. Subjects were instructed not to move their head and to look straight ahead at the screen and were asked to fixate the white fixation point before the start of the next trial.

Each run (300 scans) was composed of 48 trials. These trials consisted of 24 property verification trials for written words and 24 property verification trials for spoken words. The experiment consisted of 8 runs and across all 8 runs each concept appeared 16 times: 8 times as written words and 8 times as spoken words. For each input-modality, every concept was combined with each property once per subject. Thus across all 8 runs a combination of a given stimulus with a given property never recurred within a specific input-modality. The properties were presented equally often in written as in spoken modality across all subjects.

2.3 Image acquisition

A 3 Tesla Philips Achieva equipped with a 32-channel head volume coil provided functional and structural images. Structural imaging sequences consisted of a T1-weighted 3D turbo-field-echo sequence (repetition time = 9.6 ms, echo time = 4.6 ms, in-plane resolution = 0.97 mm, slice thickness = 1.2 mm). Functional images were obtained using T2* echoplanar images comprising 36 transverse slices (repetition time = 2 s, echo time = 30 ms, voxel size 2.75 x 2.75 x 3.75 mm³, slice thickness = 3.75 mm, and sensitivity encoding (SENSE) factor = 2), with the field of view (FOV) (220 x 220 x 135 mm³) covering the entire brain. Each run was preceded by 4 dummy scans to allow for saturation of the BOLD signal.

2.4 Image preprocessing

The images were analyzed with Statistical Parametric Mapping SPM8 (Wellcome Trust Centre for Neuroimaging, University College London, UK). First the data were realigned and resliced and a mean functional image was created. Scans were corrected for slice acquisition time. Next the structural image was co-registered with the mean functional image and segmented in gray matter, white matter and cerebrospinal fluid. The segmented images were saved in the native space and the normalized images were saved without modulation. The segmentation procedure also gives the transformation (linear and non-linear) from subjects space to the Montreal Neurological Institute (MNI) space. Subsequently functional images were warped to MNI using this transformation and resliced to a voxel size of 3 x 3 x 3 mm³ (Friston et al., 1995). The transformation was also applied to the structural image which was resliced to a voxel size of

$1 \times 1 \times 1 \text{ mm}^3$. Structural image, mean functional image and functional images were smoothed using a Gaussian filter with a kernel size of $5 \times 5 \times 7 \text{ mm}^3$. We used a standard SPM8 modelling to remove covariates of no interest (motion regressors and low-frequency trends).

2.5 Definition of the left perirhinal VOI

Our primary VOI was imported from Bruffaerts et al. (2013b) (Fig. 3A) but we also verified the results in the more restricted perirhinal VOI drawn manually (Fig. 3B).

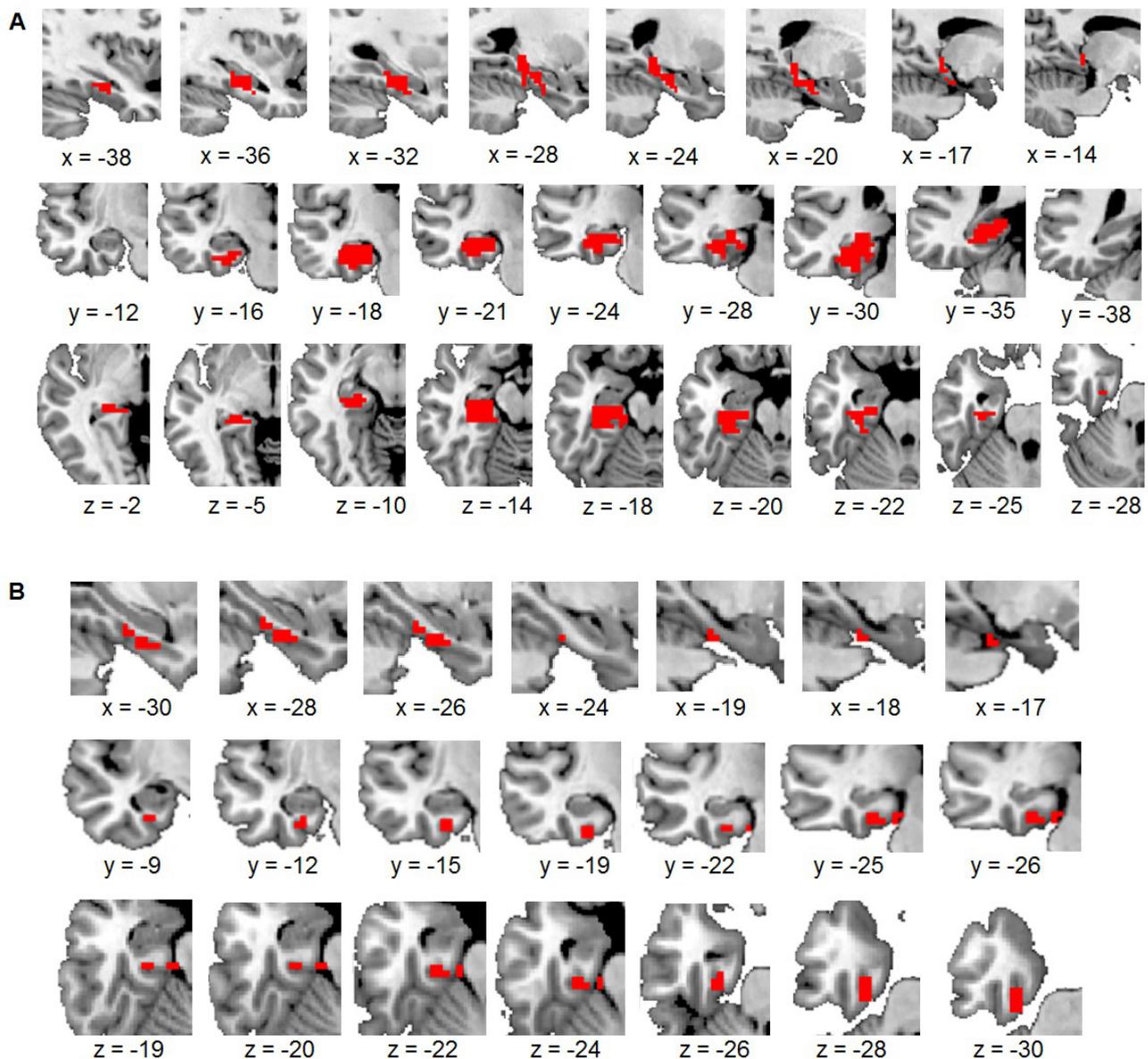


Fig. 3 (A) Left anteromedial temporal VOI imported from Bruffaerts et al. (2013b). (B) Left perirhinal VOI drawn manually based on anatomical landmarks following Insausti et al. (1998) and Kivisaari et al. (2013).

The perirhinal VOI was imported from Bruffaerts et al. (2013b). Bruffaerts et al. (2013b) defined this VOI based on a univariate contrast between property verification for written words and pictures minus a low-level control condition. The property verification condition was nearly identical to the one described above, apart from some differences in timing due to the use of spoken modality in our study. In the low-level control condition, Bruffaerts et al. (2013b)

replaced the written words and pictures by consonant letter strings and scrambled pictures, respectively, which were followed by a simple question, i.e. whether the stimulus was derived from a word or a picture (50% yes responses). This contrast yielded an extensive ventral occipitotemporal zone of activation, along with other activations. The left occipitotemporal activation zone showed a semantic similarity effect for written words. In an effort to localize the anatomical source for this effect with higher precision, Bruffaerts et al. (2013b) divided the occipitotemporal zone in 5 equally sized portions along a rostrocaudal axis and subsequently tested for the semantic similarity effect in each of these portions. Only the second most anterior portion (size = 162 voxels) showed a significant semantic similarity effect for words. We show this portion in detail in Fig. 3A. Its mathematical centre is at $x = -26$, $y = -25$, and $z = -14$, its most anterior border at $y = -15$ mm and its most posterior border at $y = -35$ mm. The VOI contains the posterior half of perirhinal cortex but also an anterior portion of the parahippocampal gyrus and extends medially into the transentorhinal and entorhinal cortex and hippocampal body (Insausti et al., 1998; Kivisaari et al., 2013) (Fig. 3A). While we refer to this as the perirhinal VOI, it would be strictly speaking more accurate to refer to it as the "perirhinal-centered" VOI.

We verified our results in a left perirhinal VOI that was defined based on anatomical landmarks (Insausti et al., 1998; Kivisaari et al., 2013). We manually delineated the left perirhinal cortex on the single subject T1 template of SPM8 based on the anatomical landmarks following the guidelines of Insausti et al. (1998) and Kivisaari et al. (2013). From this VOI we excluded voxels that fell outside the mask obtained in the univariate SPM analysis in any of the subjects. The most anterior border of this perirhinal VOI is at $y = -9$ mm and its most posterior border at $y = -26$ mm. We show this VOI (size = 24 voxels) in detail in Fig. 3B. This same 24-voxel VOI was used in each of the subjects.

2.6 Semantic similarity analysis

The MVPA analysis procedures have been described before (Bruffaerts et al., 2013a,b). It was based on the fMRI responses derived from the General Linear Model analysis using SPM8 by calculating the area under the curve of the BOLD response between 2 and 8 s after the start of the stimulus within every voxel. For each trial, a vector was constructed in a higher-dimensional space, with the activity levels in each voxel of the VOI 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 VOI. The matrix containing the pairwise cosine similarity between every two trials will be called the result matrix.

We evaluated the effect of semantic similarity on the fMRI activity pattern in two ways: In a representational similarity analysis (RSA) (Kriegeskorte et al., 2008; Bruffaerts et al., 2013b), we examined the cosine similarity between the cosine similarity matrix based on the fMRI activity patterns (called the "fMRI cossimilarity matrix") and the semantic cossimilarity matrix (De Deyne et al., 2008). We did so for the fMRI cossimilarity matrix for written and

spoken words pooled, for written words only, for spoken words only, and for exclusively cross-modal written-spoken word pairs. In a second approach, we evaluated whether pairs belonging to a same semantic subcategory (as determined based on the hierarchical clustering analysis, Fig. 1) evoked response patterns that were more similar than what is obtained with random permutation labelling. Third, we also evaluated the similarity of fMRI responses to repetitions of a same item.

2.6.1 Representational similarity analysis

We calculated the cosine similarity between semantic cossimilarity matrix and the entity-by-entity matrix of cosine similarities based on fMRI activity patterns in the perirhinal VOI (Fig. 3A) ("fMRI cossimilarity matrix").

The fMRI cossimilarity matrix was constructed as follows: the cosine similarity between two entities i and j was obtained by averaging the cosine similarity between trials of entity i and trials of entity j within a subject and by averaging this value across subjects. We have calculated four different types of such matrices depending on the input-modality of the trial (written or spoken words): the first was generated by taking all trials of entities i and j into account and we refer to this as "written and spoken words pooled". The second and the third matrices were generated by taking only trials of a same input-modality into account, i.e. trials of entity i and trials of entity j were presented in the same input-modality. The resulting matrices were called "written words" and "spoken words" matrices, respectively. The fourth matrix was generated by using only pairs of trials in which the trials of entity i and trials of entity j belonged to a different input-modality. This matrix was called "cross-modal".

We determined the cosine similarity between the group-averaged fMRI cossimilarity matrix and the semantic cossimilarity matrix. Note that we first average the fMRI cossimilarity matrix across subjects and then determine the cosine similarity with the semantic cossimilarity matrix. The significance of results was determined by performing a 10,000 random permutation labelling of the matrix based on fMRI data. We assigned a random label to each trial 10,000 times and recalculated the cosine similarity. By comparing the true cosine similarity with the 10,000 possible results obtained by chance, we could determine the significance of the result. We used a one-tailed statistical threshold of $P < 0.05$ (Bruffaerts et al., 2013b).

In order to evaluate the normality of the distribution across subjects, we determined the cosine similarity between each individual's fMRI cossimilarity matrix and the semantic cossimilarity matrix. We evaluated whether the distribution of these cosine similarities deviated from normality (Shapiro-Wilk) and also whether the distribution differed from the null distribution (Student's t test). Next, we computed for each individual a standardized position within the distribution by calculating the P value for the cosine similarity between the individual's fMRI cossimilarity matrix and the semantic cossimilarity matrix (based on 10,000 random permutation labellings, see above) and transforming the P value to a Z score. We examined whether the distribution of the Z scores deviated from normality (Shapiro-Wilk) and whether the distribution differed from the null distribution (Student's t test).

We also directly compared the cosine similarity between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written versus auditory input using a non-parametric

test to determine the significance of the difference in RSA values (= cosine similarity between semantic cossimilarity and fMRI cossimilarity) for written versus spoken words. The non-parametric test consisted of randomly (10,000 randomizations) mixing written and spoken words. Concretely, to obtain a vector (with a length equal to the number of entities) with 0 and 1, we randomly (binomial with equal probability) assigned 0 or 1 to each of the 24 entities and hence obtained a vector which can have an arbitrary number of zeros or ones. There were no constraints used during this randomization. If both entities within a pair had a value of 0, we maintained the value of the corresponding cell within the fMRI cossimilarity matrix for the written and the spoken words. When they both had a value = 1, the values were swapped between the two matrices. When their value differed, the value in the cell was replaced in both matrices by the crossmodal fMRI cossimilarity of this pair of entities. By calculating the difference in RSA values (cosine similarity of each of these two newly constructed matrices with the semantic cossimilarity matrix), we could obtain a distribution of the difference when mixing written and spoken words in a random way and we could compare the true difference against this distribution.

2.6.2 Effect of semantic subcategory

We determined whether the cosine similarity between pairs of entities belonging to the same semantic subcategory (according to the hierarchical clustering analysis, Fig. 1) was higher than that between pairs based on random permutation labelling (Bruffaerts et al., 2013b). We selected from the result matrix the cosine similarity related to pairs of trials containing entities belonging to a same semantic subcategory. We then averaged the cosine similarity over all pair of trials of the same semantic subcategory within each subject and over subjects. An average cosine similarity for each of the 6 semantic subcategories was obtained. Next we averaged these values over all 6 semantic subcategories and the result is denoted by ACS. As for the RSA, we did this for written and spoken words pooled, for written words only, for spoken words only and for cross-modal pairs.

The significance of results was determined by performing a 10,000 random permutation labelling. We assigned a random label to each trial 10,000 times and recalculated the ACS. Next we compared the true cosine similarity with the 10,000 possible results obtained by chance. We used a one-tailed statistical threshold of $P < 0.05$ (Bruffaerts et al., 2013b).

2.7 Repetition analysis

In the repetition analysis, we determined whether the cosine similarity between pairs of trials related to the same entity was higher than that between pairs based on random permutation labelling (Bruffaerts et al., 2013b). We selected from the result matrix the cosine similarity related to pairs of trials containing the same entity. We then averaged the cosine similarity over all pair of trials related to the same entity within each subject and over subjects. An average cosine similarity for each of the 24 entities was obtained. Next we averaged these values over all the 24 entities and the result is denoted by ACS.

The significance of results was determined by performing a 10,000 random permutation

labelling. We assigned a random label to each trial 10,000 times and recalculated the ACS. Next we compared the true cosine similarity with the 10,000 possible results obtained by chance. We used a one-tailed statistical threshold of $p < 0.05$ (Bruffaerts et al., 2013b).

2.8 Evaluation of orthographic variables on perirhinal VOI

In order to evaluate the effect of orthographic variables upon the perirhinal activity pattern, we performed an RSA between the orthographic distance matrix (based on Levenshtein distances) and the fMRI cosine distance matrix in left perirhinal cortex. The fMRI cosine distance matrix is calculated by subtracting the cosine similarities from 1 in each cell. Otherwise, the procedure was exactly identical to that applied for the RSA with the semantic cossimilarity matrix.

3 Results

3.1 Behavioral analysis

Reaction times were calculated from the onset of the question. Reaction times differed significantly between conditions presented in the written input-modality (mean = 1,81 sec, S.D. = 0.47) and those presented in the spoken input-modality (mean = 2,27 sec, S.D. = 0.43) (Student's $t(20) = 3,2168$, $p < 0.001$).

We conducted a 6x8 repeated measures ANOVA with two factors (subcategory: 6 levels, and property: 8 levels) and with reaction times as outcome. The main effect of subcategory was significant ($F(5,95) = 7.35$, $p = 0.00001$). The main effect of property was also significant ($F(7,133) = 32.29$, $p = 0.0000$). There was a significant interaction between subcategory and property ($F(36,665) = 3.99$, $p = 0.0000$). According to Tukey posthoc testing, subjects responded more slowly to fish and amphibians compared to the other subcategories. They responded more rapidly for the property question "does it fly?" than for the other properties. They responded more slowly for the property "is it exotic?", especially in combination with the subcategories marine animals, fish, and amphibians. They also responded more slowly for the property "is it a mammal", especially in combination with the subcategory marine animals.

3.2 MVPA of left anteromedial temporal VOI imported from Bruffaerts et al. (2013b)

In view of the well-known fMRI susceptibility artefact in anterior temporal cortex, we calculated the temporal signal-to-noise ratio (TSNR) by dividing the mean of the residual time series by its standard deviation (Murphy et al. (2007)). We provide an image of TSNR obtained in this experiment in Fig. 4. The TSNR in our VOI ranged from 133 to 183.

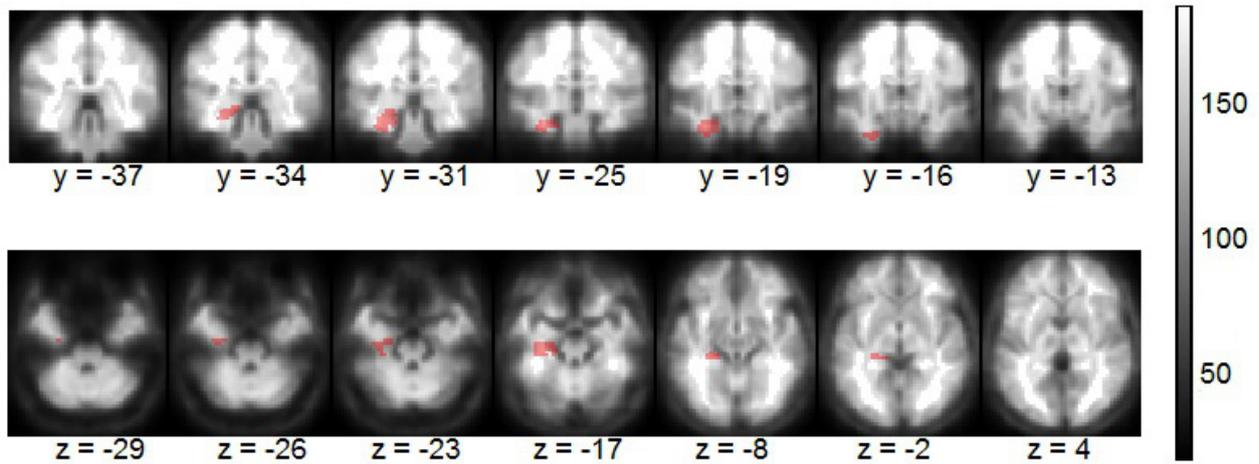


Fig. 4 Temporal signal-to-noise-ratio (TSNR) averaged over all 20 subjects. The TSNR is calculated by dividing the mean of the time series by its standard deviation (Murphy et al., 2007).

In the left anteromedial temporal volume imported from Bruffaerts et al. (2013b), RSA revealed a significant correlation between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written words (cosine similarity = 0.4934, $P = 0.0033$) (Fig. 5A). This replicates the effect originally reported by Bruffaerts et al. (2013b). No effects were seen for spoken words (cosine similarity = 0.4685, $P = 0.347$), for written and spoken words pooled (cosine similarity = 0.1571 $P = 0.3248$), or cross-modally (cosine similarity = -0.3920, $P = 0.8373$) (Fig. 5B-D).

Table 1 Left anteromedial temporal VOI imported from Bruffaerts et al. (2013b) (Fig. 3A). Representation similarity analysis, semantic similarity and repetition effect. Values which reached significance at uncorr. $P < 0.05$ are in bold.

Written and spoken words pooled		
RSA	CS = 0.1571	P = 0.3248
Similarity within semantic subcategories	ACS = 0.0015	P = 0.2448
Repetition effect	ACS = 0.0019	P = 0.2448
Written words		
RSA	CS = 0.4934	P = 0.0033
Similarity within semantic subcategories	ACS = 0.0109	P = 0.005
Repetition effect	ACS = 0.0094	P = 0.1969
Spoken words		
RSA	CS = 0.4685	P = 0.3431
Similarity within semantic subcategories	ACS = 0.0072	P = 0.4854
Repetition effect	ACS = 0.0098	P = 0.2242
Cross-modal effect		
RSA	CS = -0.3920	P = 0.8373
Similarity within semantic subcategories	ACS = -0.0059	P = 0.8304
Repetition effect	ACS = -0.0048	P = 0.4908

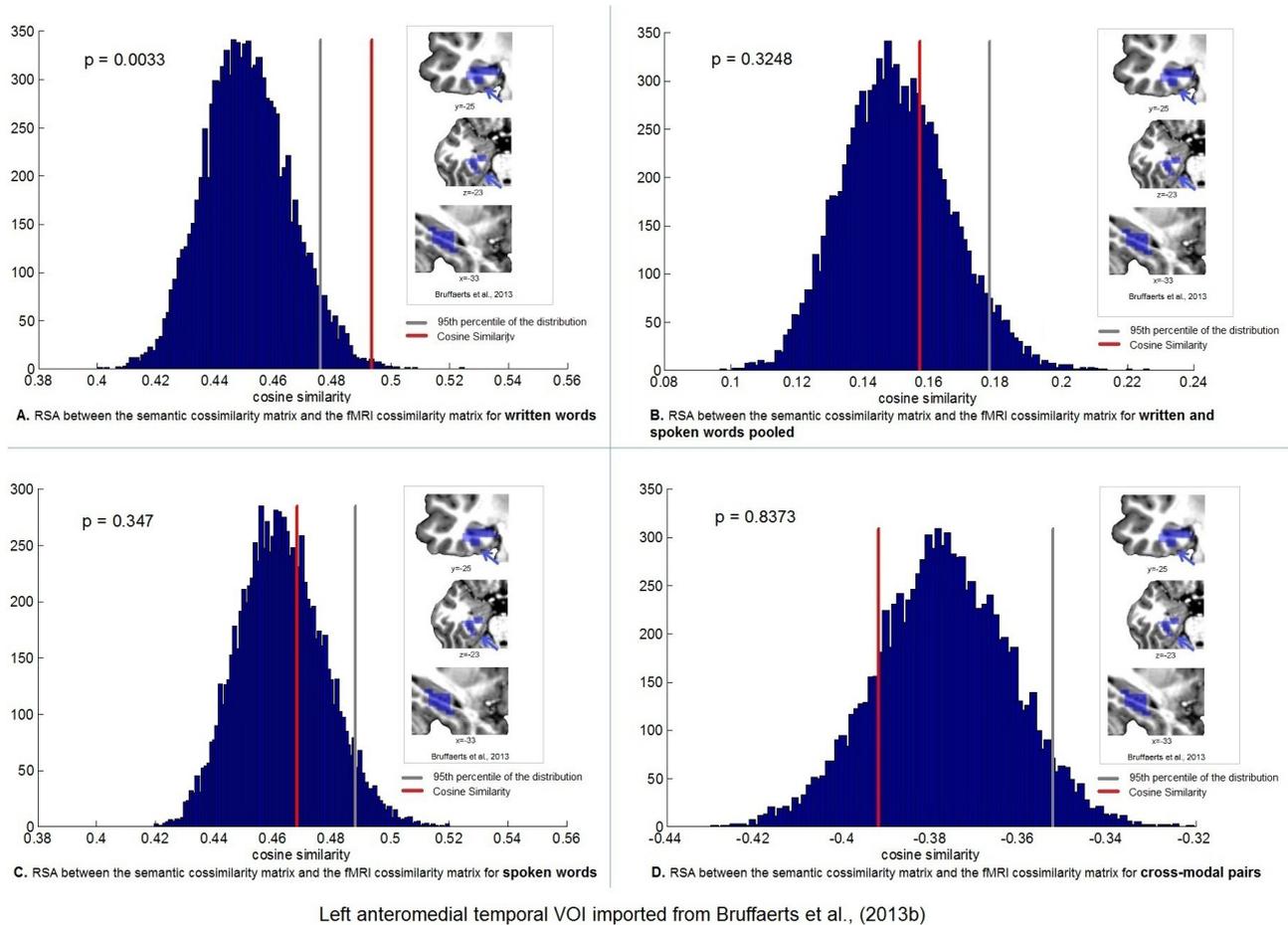
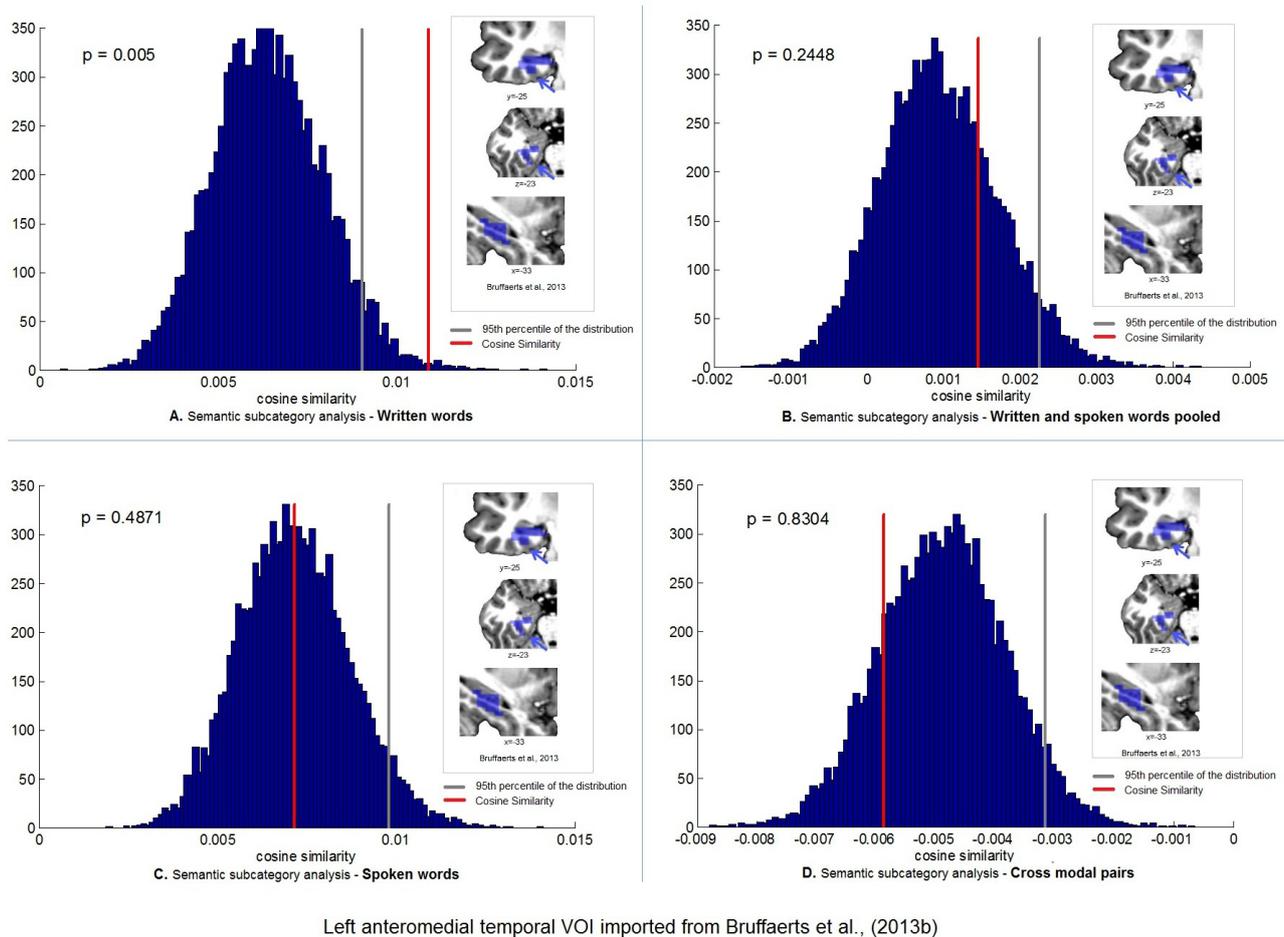


Fig. 5 Left anteromedial temporal VOI imported from Bruffaerts et al. (2013b). Probability distributions for the representational similarity analysis (RSA). The red line indicates the cosine similarity between the similarity matrix based on behavioral data (Deyne and Storms, 2008) and the similarity matrix based on the fMRI data derived from the response patterns within left anteromedial temporal cortex (Bruffaerts et al., 2013b). The gray line indicates the 95th percentile of the distribution. 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. (A) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written words. (B) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written and spoken words pooled. (C) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for spoken words. (D) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for cross-modal pairs.

We also directly compared the cosine similarity between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written versus spoken words. No significant difference was observed ($P = 0.34$).

Within the written modality we assessed the distribution of the individuals' CS values. The distribution did not significantly deviate from normality (Shapiro-Wilk test, $P = 0.0696$) and significantly differed from the null distribution (Student's $t = 5.1516$, $df = 19$, $P < 0.00005$). Likewise, the distribution of the Z scores did not significantly deviate from normality (Shapiro-Wilk test, $P > 0.23$) and significantly differed from the null distribution (Student's $t = 4.1592$, $df = 19$, $P < 0.0003$).



Left anteromedial temporal VOI imported from Bruffaerts et al., (2013b)

Fig. 6 Left anteromedial temporal VOI imported from Bruffaerts et al. (2013b). Probability distributions for the effect of semantic subcategory. The red line indicates the average cosine similarity (ACS) of the left anteromedial temporal VOI ((Bruffaerts et al., 2013b) between entities belonging to a same cluster, presented as (A) written words, (B) written and spoken words pooled, (C) spoken words, and (D) cross-modal pairs. X axis: cosine similarity averaged over the group of subjects. Y axis: absolute frequency of a given average cosine similarity value across a total of 10,000 random permutation labellings. Gray line: 95th percentile of the distribution.

In the semantic subcategory analysis, ACS was significantly higher for pairs of written words originating from a same semantic subcategory than for pairs derived from random permutation labelling (ACS = 0.0109 $P = 0.005$) (Table 1). No effects were seen for spoken words (ACS = 0.0072, $P = 0.4871$), for written and spoken words pooled (ACS = 0.0015, $P = 0.2448$) and for cross-modality (ACS = -0.0059, $P = 0.8304$) (Fig. 6).

In the repetition analysis, no effects were seen for either of the two input modalities (Table 1).

3.3 MVPA of manually drawn left perirhinal VOI

In the manually drawn left perirhinal volume, RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix was near significance for written words ($P = 0.0555$) (Table 2) (Fig. 7).

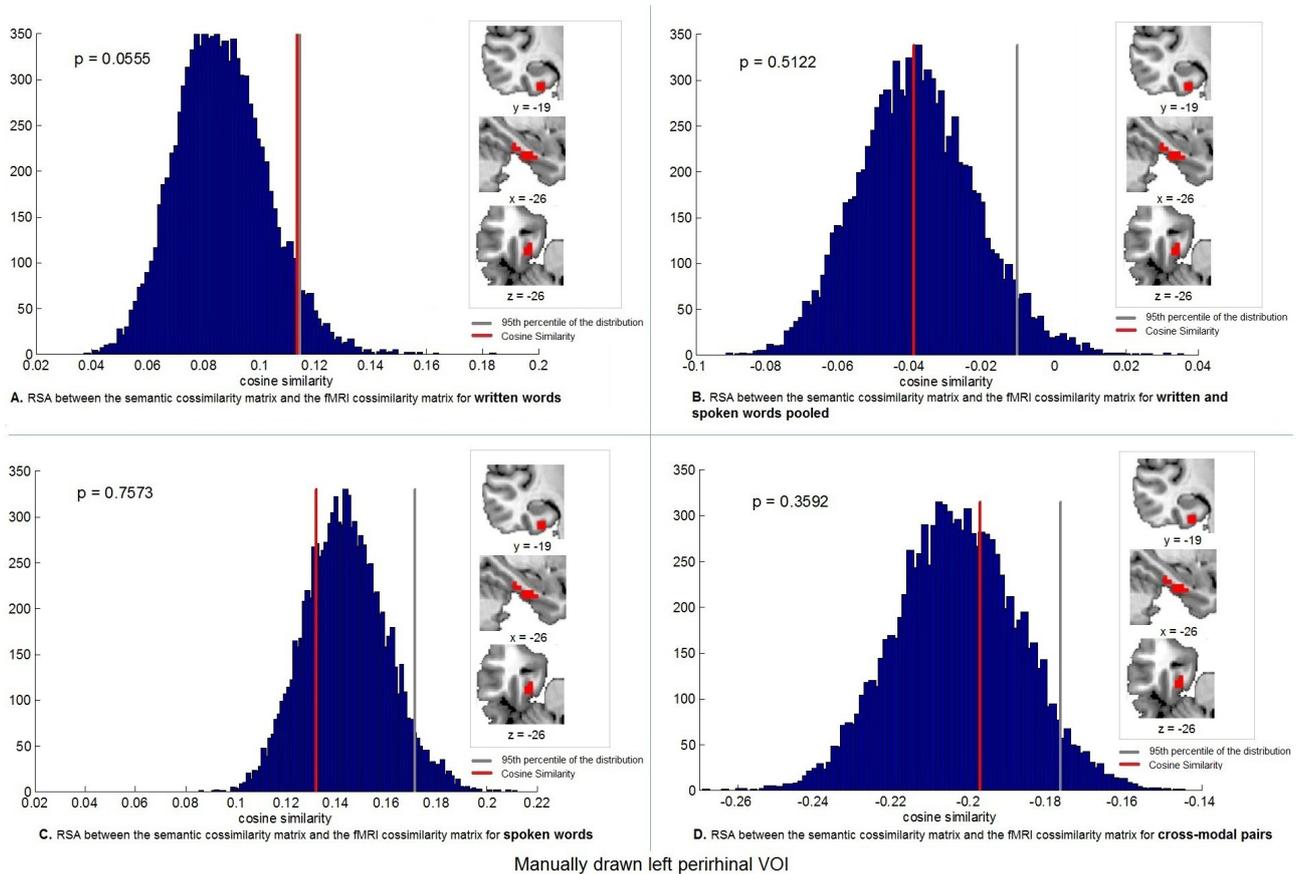


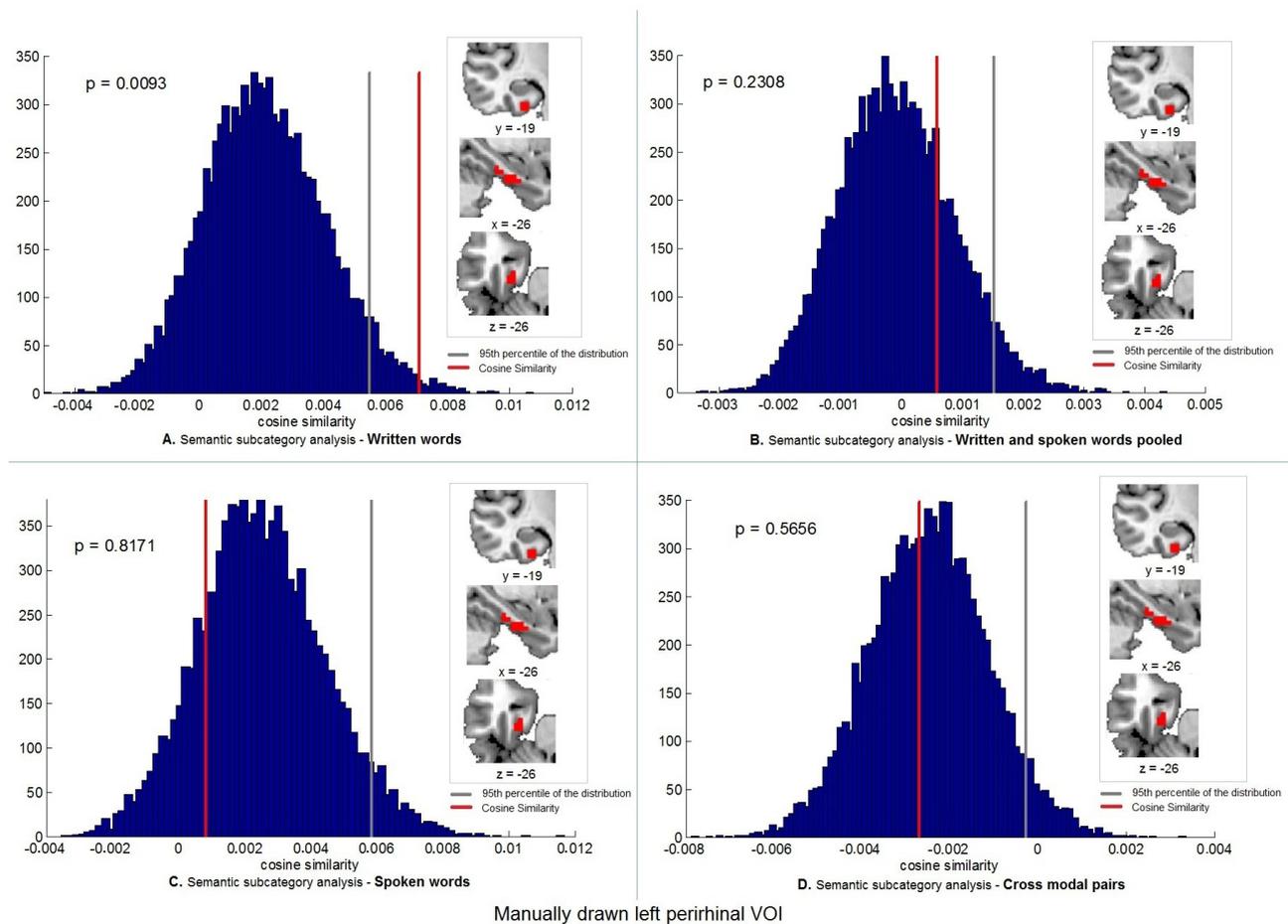
Fig. 7 Manually drawn left perirhinal VOI: Probability distributions for the representational similarity analysis (RSA). Conventions see Fig. 5. The red line indicates the cosine similarity between the similarity matrix based on behavioral data (Deyne and Storms, 2008) and the similarity matrix based on the fMRI data derived from the response patterns within left perirhinal cortex with words as input modality. (A) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written words. (B) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written and spoken words pooled. (C) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for spoken words. (D) RSA between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for cross-modal pairs.

In the semantic subcategory analysis, ACS was significantly higher for pairs of written words originating from a same semantic subcategory than for pairs derived from random permutation labelling (ACS = 0.0071 $P = 0.0074$) (Table 2) (Fig. 8). No effects were seen for spoken words (ACS = 0.0008, $P = 0.8171$), for written and spoken words pooled (ACS = 0.0005, $P = 0.2308$) and for cross-modality (ACS = -0.0027, $P = 0.5656$).

In the repetition analysis, ACS was significantly higher for pairs of written words related to the same entity than for pairs derived from random permutation labelling (ACS = 0.0128 $P = 0.0081$) and for written and spoken words related to the same entity than for pairs derived from random permutation labelling (ACS = 0.0056 $P = 0.0052$) (Table 2). No effects were seen for spoken words (ACS = 0.0079, $P = 0.1092$) and for cross-modal pairs (ACS = 0.0015, $P = 0.0828$).

Table 2 Manually delineated perirhinal VOI (Fig. 3B). Representation similarity analysis, semantic similarity and repetition effect. Values which reached significance at uncorr. $P < 0.05$ are in bold.

Written and spoken words pooled		
RSA	CS = -0.0393	P = 0.5122
Similarity within semantic subcategories	ACS = 0.0005	P = 0.2308
Repetition effect	ACS = 0.0056	P = 0.0052
Written words		
RSA	CS = 0.1135	P = 0.0555
Similarity within semantic subcategories	ACS = 0.0071	P = 0.0074
Repetition effect	ACS = 0.0128	P = 0.0081
Spoken words		
RSA	CS = 0.1321	P = 0.7498
Similarity within semantic subcategories	ACS = 0.0008	P = 0.809
Repetition effect	ACS = 0.0079	P = 0.1092
Cross-modal effect		
RSA	CS = -0.1972	P = 0.3592
Similarity within semantic subcategories	ACS = -0.0027	P = 0.5656
Repetition effect	ACS = 0.0015	P = 0.0828

**Fig. 8** Manually drawn left perirhinal VOI. Probability distributions for the effect of semantic subcategory. Conventions see Fig. 6. The red line indicates the average cosine similarity (ACS) between entities belonging to a same cluster, presented as (A) written words, (B) written and spoken words pooled, (C) spoken words, and (D) cross-modal pairs. X axis: cosine similarity averaged over the group of subjects. Y axis: absolute frequency of a given average cosine similarity value across a total of 10,000 random permutation labellings. Gray line: 95th percentile of the distribution.

3.4 Evaluation of effects of orthographic-lexical variables

The cosine similarity between the orthographic distance matrix and the fMRI cosine distance matrix was not significant ($P > 0.93$).

4 Discussion

In left anteromedial temporal cortex, the similarity between fMRI response patterns to written words reflects the similarity in meaning between written words (Figs. 5A and 6A). This replicates our previous findings (Bruffaerts et al., 2013b). However, contrary to our predictions, for spoken words, we did not obtain a semantic similarity effect in this region (Table 1 and Figs. 5C and 6C). The present data differ from what we had expected based on a hypothetical role of this anteromedial temporal region as an "amodal hub" for semantic processing.

4.1 Alternative explanations and limitations

For the majority of words, the word form has no relationship with the meaning of the word, or, put otherwise, with the pairwise similarities derived from the concept-feature matrix. Since we only found significant effects for the written word modality, we thoroughly evaluated any potential effect of orthographic variables. There was no relationship between the semantic cossimilarity matrix and word length, orthographic neighbourhood size or orthographic distance. Neither did we find any relationship between the orthographic distance matrix and the perirhinal activity patterns. For these reasons, an explanation in terms of orthographic string variables can be dismissed.

We also examined whether any of the classical word parameters, word frequency, word familiarity, age-of-acquisition and imageability, was reflected by the semantic cossimilarity matrix. When the semantic distance between two entities was higher, they were also more likely to differ more in these parameters. In principle, these word characteristics are shared between the written and the spoken modality and are unlikely as an explanation for the semantic similarity effect in perirhinal cortex as we obtained this effect for the written modality only. A link between these variables and the meaningful content of the words is well-known and does not exclude an interpretation of our findings in terms of semantic word content at all.

Our MVPA approach is based on a prior volume of interest defined in an independent dataset. The VOI imported from Bruffaerts et al. (2013b) contained the posterior half of perirhinal cortex, and also extended into the rostral portion of the parahippocampal gyrus and medially into entorhinal cortex and the hippocampus (Insausti et al., 1998; Kivisaari et al., 2013) (Fig. 3). To address this issue we manually delineated the perirhinal cortex based on anatomical landmarks (Fig. 3B). Overall, this confirmed the findings (Table 2 and Figs. 7A and 8A).

We found significant effects for written but not for spoken word input. However, the direct contrast of the cosine similarity between the semantic cossimilarity matrix and the fMRI cossimilarity matrix for written versus spoken words did not yield a significant difference between

the two input-modalities. We therefore have to interpret the relative exclusivity for written word input cautiously.

4.2 Perirhinal cortex and associative coding between objects

Perirhinal cortex, entorhinal cortex and parahippocampal gyrus (Fig. 3) have been grouped as the "parahippocampal component of the paralimbic belt" by Mesulam (1998). The entorhinal cortex has a peri-allocortical architecture while the perirhinal cortex architecture is peri-isocortical (Mesulam, 2000). Which neurophysiological mechanisms could underly the correlation between the semantic cossimilarity matrix and the fMRI cossimilarity matrix in this region?

A possible mechanism could be put forth in analogy with the pair coding neurons found in perirhinal cortex in nonhuman primates (NHP) (Yoshida et al., 2003; Naya et al., 2003; Hirabayashi et al., 2013; Suzuki and Naya, 2014). In the paired-associates learning paradigm (Yoshida et al., 2003; Naya et al., 2003; Hirabayashi et al., 2013; Suzuki and Naya, 2014), NHP first learn the association between two stimuli which form a yoked pair in a list of, e.g. 10 pairs. When a visual stimulus is presented that belongs to a learned pair, a pair-coding neuron is defined as a neuron where the firing rate in response to the visual stimulus correlates with the firing rate of the response to its paired associate, even when the stimuli are shown separately from each other (Naya et al., 2003). The percentage pair-coding neurons dramatically increases from the inferotemporal region TE (4.9% of the cue-selective neurons) to perirhinal cortex (BA36) (33%) (Naya et al., 2003; Hirabayashi et al., 2013). The median paired associates correlation coefficient is also higher in perirhinal cortex (0.51) than in TE (0.14) (Yoshida et al., 2003). According to Naya et al. (2003), selective convergence in perirhinal cortex is the neuronal basis of cortical cell assembly for "semantic-like memory, which is the partner to episodic-like memory that may require more distributed networks in the hippocampus" (Naya et al., 2003). Early ablation studies in NHP also support the notion that perirhinal cortex is involved in associative coding during perception of objects and semantic memory (Murray and Bussey, 1999).

We conjecture that the neurophysiological "pair-coding" effect may underly the cosine similarity of fMRI response patterns between words that are closer in meaning. Obviously, one has to be careful drawing analogies between single-neuron recording findings in NHP and human fMRI. Conceptually, the paired associates learning paradigm has a well-controlled learning phase starting with novel stimuli while the concept-feature matrix integrates learned experiences from early childhood onward. Furthermore, the spatial resolution of single neuron and multi-unit recordings in perirhinal cortex is many orders of magnitude higher than that of our MRI method. Third, the exact homology between perirhinal cortex in NHP and our region would need to be established through direct inter-species comparison. The perirhinal, entorhinal and transentorhinal cortex have expanded in humans compared to monkeys (Mesulam, 1998). Finally, the link between processing written words and processing visual objects remains hypothetical, also from an evolutionary viewpoint. We (Vandenberghe et al., 1996) and others (Mesulam, 2000) have previously speculated that in the ventral occipitotemporal cortex, visual word processing developed on the foundations of a pre-existing neuroanatomical path-

ways for visual object recognition. We conjecture that the correlation in neuronal responses to learned associates of a pair of visual objects in perirhinal cortex (Suzuki and Naya, 2014) forms the closest single neuron electrode analog to date of the cosine similarity between the semantic cossimilarity matrix and the fMRI cossimilarity matrix.

4.3 Extracting meaning from printed words

Our findings demonstrate the role of left perirhinal cortex in coding the similarity between entities represented as *written words*, rather than visual objects. Previous fMRI studies have revealed effects of conceptual priming with written words in perirhinal cortex (O’Kane et al., 2005; Voss et al., 2009; Wang et al., 2010). Furthermore, medial temporal lobe lesions extending into perirhinal cortex virtually abolish conceptual priming (Wang et al., 2010). Perirhinal cortex is also involved in explicit semantic tasks with written words. For instance, during a typicality judgment for written words and pictures, RSA revealed a crossmodal effect both for written words and for pictures in left ventral temporal cortex, encompassing among other areas the perirhinal cortex (-33, -29, -23) (Fairhall and Caramazza, 2013). In another study, an associative-semantic judgment with written words activated ventromedial temporal cortex comprising among other areas perirhinal cortex (Visser et al., 2010a). Further evidence that perirhinal cortex plays a role in written word processing comes from invasive electrophysiological studies in humans (Nobre and McCarthy, 1995). Local field potential recordings, current source density analysis and single neuron recording in human perirhinal cortex indicate that during judgement of real-life size, familiarity or concreteness as well as during overt word recognition, written words evoke distinct effects in perirhinal cortex (Halgren et al., 2006; Chan et al., 2011).

4.4 No detectable semantic similarity effect for spoken word modality

A priori we had expected that left perirhinal cortex would also exhibit a semantic similarity effect for the spoken word modality, in agreement with a putative role in amodal semantic processing. As mentioned above, in the Fairhall and Caramazza (2013) study RSA revealed semantic similarity effects for both written words and pictures in ventral temporal cortex, indicative for commonality at least between visual input modalities. According to one prevailing model, perirhinal cortex functions as a polymodal convergence zone where features from multiple input channels are integrated (Murray and Bussey, 1999). We could not confirm this hypothesis: the left anteromedial temporal cortex exhibited a semantic similarity effects for written words, not for spoken words (Figs. 5A, C and 6A,C). One always has to be careful interpreting negative findings. Sensitivity to detect effects for spoken input could have been lower than for visual input, e.g. if auditory input is projected to different cortical laminae than visual input and if this entails differences in the sensitivity of the BOLD signal. We however put forward that the effect for written words in the absence of an effect for auditory stimuli may reflect a specialization of perirhinal cortex for visual input.

Perirhinal cortex is a predominantly visual area at the top of the visual processing pathway. 65% of afferents to PRH come from the inferotemporal unimodal visual association areas TE

and TEO (Insausti et al., 1987), with also a connection from BA36 to the upper bank of the rostral end of STS, a polymodal area (Saleem et al., 2000). This connectivity pattern differs drastically from parahippocampal regions such as PF, which is heavily connected with retrosplenial and posterior parietal cortex (Suzuki and Amaral, 1994a,b; Libby et al., 2012). Studies that examined integration between modalities mainly focussed on the integration between visual and tactile (Holdstock et al., 2009) and visual and flavor (Parker and Gaffan, 1998) rather than auditory input.

In two out of nine cases who underwent deep electrode recordings, Chan et al. (2011) recorded also for spoken and not only for written words and found similar responses to the spoken as to the written words. Likewise, using magnetoencephalography, Marinkovic et al. (2003) compared real-life size judgment based on visual versus auditory words. While the initial stages differed, convergence was seen in the anterior temporal pole. These findings clearly demonstrate that the anterior temporal cortex contains regions that respond to both written and auditory words. The absence of an effect for auditory word input in our study may be related to the difference in measurement method, fMRI versus electrophysiological measures. Alternatively, our anteromedial temporal VOI may not contain the sites of origin of the electrophysiological effects that are shared between auditory and visual word input (Marinkovic et al., 2003; Chan et al., 2011). The recording site in the Chan et al. (2011) cases appears to be slightly more lateral, at the medial border of anterior occipitotemporal sulcus. Possibly, the recording sites fall outside the volume of interest examined in the current study. This idea is also supported by a PET study of listening to narrative speech for auditory or written input modality (Spitsyna et al., 2006) as well as by a meta-analysis comparing semantic processing of auditory versus visual materials in the anterior temporal pole (Visser et al., 2010a). Like the electrophysiological studies (Marinkovic et al., 2003; Chan et al., 2011), the PET study (Spitsyna et al., 2006) showed convergence between auditory and written word modality in the anterior temporal pole laterally to the perirhinal cortex both for the written and the spoken modality ("lateral TP"). Because of the lateral distribution, Spitsyna et al. (2006) added a contrast between written minus auditory rotated stimuli and false fonts. This revealed more ventral and medial activation for the written modality specifically (Nobre et al., 1994). In another study comparing auditory versus visual semantic tasks (Visser and Lambon Ralph, 2011), the 5-mm radius spherical volumes of interest that were activated in common for auditory and visual input modality were more ventral than our VOI (Fig. 3). This was also the case for the anterior inferior temporal activation obtained for the conjunction of visual and auditory input (-39, -18, -30 versus Fig. 3) (Visser and Lambon Ralph, 2011).

Taken together, the electrophysiological and PET studies and our findings lead us to hypothesize that regions in anterior temporal cortex adjacent to our anteromedial temporal VOI, e.g. more ventrally, laterally or rostrally, show semantic effect for both written and auditory words. In the current dataset, we have not evaluated these adjacent regions because of the relatively low TSNR in these regions, a well-known MRI susceptibility issue (Fig 4). Based on their connectivity pattern (Morán et al., 1987) the more lateral and rostral regions may be better candidates as amodal hubs than perirhinal cortex itself.

4.5 Visual word processing model

According to NHP studies, more posterior regions in the visual processing pathway such as the inferotemporal areas TE and TEO are involved in object identification at the single object level while perirhinal cortex deals with associative coding between visual objects (Orban et al., 2014; Suzuki and Naya, 2014). We can now discern a similar progression from MVPA studies in humans. Posterior fusiform cortex has been mainly implicated in identification of objects ("the structural description system") (Vandenbulcke et al., 2006; Bruffaerts et al., 2013a). To the left, a posterior fusiform area called the visual word form area (Cohen et al., 2000) shows equally strong responses when subjects consciously identify pictures or written words (Van Doren et al., 2010; Seghier and Price, 2011; Vogel et al., 2012). More anteriorly in the visual object processing pathway, in left perirhinal cortex, response patterns are influenced by semantic distance between entities and semantic subcategories into which the animate entities can be divided (Bruffaerts et al., 2013b). In our view these findings fit with the hypothesis that visual word processing is partly built on the foundations of the hierarchical visual object processing pathway in the left hemisphere (Vandenberghe et al., 1996; Mesulam, 1998). According to this visual word processing model, the ventral occipitotemporal posterior-anterior progression from word identification to word association is similar to progression proposed by Suzuki and Naya (2014) for visual objects, from object identification to associative coding between objects. As far as our data allow us to tell, at the left perirhinal stage there is convergence between written word and visual object modality (Fairhall and Caramazza, 2013) but not between written and auditory input modality. Within the "Lichtheim-2" dual-pathway framework (Ueno et al., 2011) this would mean that apart from a dorsal and a ventral, lateral temporal auditory word processing pathway, a third ventromedial temporal pathway exists for visual word input. Via this pathway orthographic-lexical representations may be directly linked to word meaning in a way that bypasses phonological retrieval or inner auditory speech. This may also explain why isolated lesions of left perirhinal cortex do not necessarily impair written word comprehension as alternative pathways are available that may be spared, e.g. via the lateral temporal pathway.

We hypothesize that true amodal convergence happens in more rostral and ventrolateral anterior temporal pole (Ueno et al., 2011; Lambon Ralph, 2014). Other candidate regions for amodal processing could be the left or right pars triangularis and opercularis, left angular gyrus (Bonner et al., 2013), left inferior temporal gyrus (Simanova et al., 2012; Fairhall and Caramazza, 2013) or middle temporal gyrus (Badre et al., 2005), or left intraparietal sulcus (Devereux et al., 2013).

4.6 Summary

There is a robust and consistent effect of semantic distance between written words on left anteromedial temporal activity patterns. We interpret this in the context of the role of left perirhinal cortex in associative learning as demonstrated in nonhuman primates (Murray and Bussey, 1999; Suzuki and Naya, 2014). Our findings can be integrated in a model where a ventral occipitotemporal written word processing pathway exists that is organized along a hierarchy from written word identification to associative coding between written words.

5 Acknowledgments

R.V. is a Senior Clinical Investigator of the Research Foundation Flanders (FWO). R.B. is a Ph.D. fellow of the Research Foundation - Flanders (FWO). Funded by FWO Grant G0660.09, KU Leuven Grant OT/08/56, OT/12/097 and Federaal Wetenschapsbeleid Belspo Inter-University Attraction Pole Grant P6/29 and P7/11. We are grateful to Astrid Van Wieringen and Eric Manders for help in preparing the auditory stimuli.

References

- Akama, H., Murphy, B., Na, L., Shimizu, Y., and Poesio, M. (2012). Decoding semantics across fMRI sessions with different stimulus modalities: a practical MVPA study. *Front Neuroinform*, 6:24.
- Baayen, H., Piepenbrock, R., and Rijn, H. V. (1993). *The CELEX Lexical Database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium.
- Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., and Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, 47(6):907–918.
- Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex*, 19(12):2767–2796.
- Bonner, M. F., Peelle, J. E., Cook, P. A., and Grossman, M. (2013). Heteromodal conceptual processing in the angular gyrus. *Neuroimage*, 71C:175–186.
- Bruffaerts, R., Dupont, P., De Grauwe, S., Peeters, R., De Deyne, S., Storms, G., and Vandenberghe, R. (2013a). Right fusiform response patterns reflect visual object identity rather than semantic similarity. *Neuroimage*, 83C:87–97.
- Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., and Vandenberghe, R. (2013b). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *J Neurosci*, 33(47):18597–18607.
- Buckner, R., Koutstaal, W., Shacter, D., and Rosen, B. (2000). Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. *Brain*, 123:620–640.
- Chan, A. M., Baker, J. M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., Cash, S. S., and Halgren, E. (2011). First-pass selectivity for semantic categories in human anteroventral temporal lobe. *J Neurosci*, 31(49):18119–18129.
- Chao, L., Haxby, J., and Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nat Neurosci*, 2:913–919.

- Chertkow, H., Bub, D., Deaodon, C., and Whitehead, V. (1997). On the status of object concepts in aphasia. *Brain Lang.*, 58:203–232.
- Clarke, A. and Tyler, L. K. (2014). Object-specific semantic coding in human perirhinal cortex. *J Neurosci*, 34(14):4766–4775.
- Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M., and Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split brain patients. *Brain*, 123:291–307.
- Connolly, A. C., Guntupalli, J. S., Gors, J., Hanke, M., Halchenko, Y. O., Wu, Y.-C., Abdi, H., and Haxby, J. V. (2012). The representation of biological classes in the human brain. *J Neurosci*, 32(8):2608–2618.
- Cree, G. S. and 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(2):163–201.
- De Deyne, S., Verheyen, S., Ameel, E., Vanpaemel, W., Dry, M. J., Voorspoels, W., and Storms, G. (2008). Exemplar by feature applicability matrices and other Dutch normative data for semantic concepts. *Behav Res Methods*, 40(4):1030–1048.
- Demb, J., Desmond, J., Wagner, A., Vaidya, C., Glover, G., and Gabrieli, J. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional MRI study of task difficulty and process specificity. *J. Neurosci.*, 15:5870–5878.
- Devereux, B. J., Clarke, A., Marouchos, A., and Tyler, L. K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *J Neurosci*, 33(48):18906–18916.
- Deyne, S. D. and Storms, G. (2008). Word associations: norms for 1,424 Dutch words in a continuous task. *Behav Res Methods*, 40:198–205.
- Fairhall, S. L. and Caramazza, A. (2013). Brain regions that represent amodal conceptual knowledge. *J Neurosci*, 33(25):10552–10558.
- Friston, K., Holmes, A., Worsley, K., Poline, J., Frith, C., Heather, J., and Frackowiak, R. (1995). Statistical parametric maps in functional imaging: a general approach. *Hum. Brain Mapp.*, 2:189–210.
- Goldberg, R., Perfetti, C., Fiez, J., and Schneider, W. (2007). Selective retrieval of abstract semantic knowledge in left prefrontal cortex. *J. Neurosci.*, 27:3790–3798.
- Halgren, E., Wang, C., Schomer, D. L., Knake, S., Marinkovic, K., Wu, J., and Ulbert, I. (2006). Processing stages underlying word recognition in the anteroventral temporal lobe. *Neuroimage*, 30(4):1401–1413.

- Hanson, S. J., Matsuka, T., and Haxby, J. V. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: is there a "face" area? *Neuroimage*, 23(1):156–166.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293(5539):2425–2430.
- Hirabayashi, T., Takeuchi, D., Tamura, K., and Miyashita, Y. (2013). Microcircuits for hierarchical elaboration of object coding across primate temporal areas. *Science*, 341(6142):191–195.
- Hodges, J., Patterson, K., Oxbury, S., and Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, 115:1783–1806.
- Holdstock, J. S., Hocking, J., Notley, P., Devlin, J. T., and Price, C. J. (2009). Integrating visual and tactile information in the perirhinal cortex. *Cereb Cortex*, 19(12):2993–3000.
- Insausti, R., Amaral, D. G., and Cowan, W. M. (1987). The entorhinal cortex of the monkey: Ii. cortical afferents. *J Comp Neurol*, 264(3):356–395.
- Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., Vainio, P., Laakso, M. P., and Pitknen, A. (1998). Mr volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *AJNR Am J Neuroradiol*, 19(4):659–671.
- Just, M. A., Cherkassky, V. L., Aryal, S., and Mitchell, T. M. (2010). A neurosemantic theory of concrete noun representation based on the underlying brain codes. *PLoS One*, 5(1):e8622.
- Kircher, T., Sass, K., Sachs, O., and Krach, S. (2009). Priming words with pictures: neural correlates of semantic associations in a cross-modal priming task using fMRI. *Hum. Brain Mapp.*, in press.
- Kivisaari, S., Probst, A., and Taylor, K. (2013). *fMRI: Basic and clinical applications*, chapter The perirhinal, entorhinal and parahippocampal cortices and hippocampus: an overview of functional anatomy and protocol for their segmentation in MR images, pages 239–267. Springer-Verlag Berlin Heidelberg.
- Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2:4.1–4.28.
- Lambon Ralph, M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philos Trans R Soc Lond B Biol Sci*, 369(1634):20120392.
- Lambon Ralph, M. A., Cipolotti, L., Manes, F., and Patterson, K. (2010). Taking both sides: do unilateral anterior temporal lobe lesions disrupt semantic memory? *Brain*, 133(11):3243–3255.

- Libby, L. A., Ekstrom, A. D., Ragland, J. D., and Ranganath, C. (2012). Differential connectivity of perirhinal and parahippocampal cortices within human hippocampal subregions revealed by high-resolution functional imaging. *J Neurosci*, 32(19):6550–6560.
- Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., and Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, 38(3):487–497.
- Mesulam, M. (1998). From sensation to cognition. *Brain*, 121:1013–1052.
- Mesulam, M. (2000). *Principles of behavioral and cognitive neurology*, chapter Behavioral neuroanatomy. Large-scale networks, association cortex, frontal syndromes, the limbic system and hemispheric specializations, pages 1–120. Oxford University Press.
- Mitchell, T. M., Shinkareva, S. V., Carlson, A., Chang, K.-M., Malave, V. L., Mason, R. A., and Just, M. A. (2008). Predicting human brain activity associated with the meanings of nouns. *Science*, 320(5880):1191–1195.
- Morán, M. A., Mufson, E. J., and Mesulam, M. M. (1987). Neural inputs into the temporopolar cortex of the rhesus monkey. *J Comp Neurol*, 256(1):88–103.
- Mur, M. (2014). What’s the difference between a tiger and a cat? from visual object to semantic concept via the perirhinal cortex. *J Neurosci*, 34(32):10462–10464.
- Mur, M., Ruff, D. A., Bodurka, J., De Weerd, P., Bandettini, P. A., and Kriegeskorte, N. (2012). Categorical, yet graded–single-image activation profiles of human category-selective cortical regions. *J Neurosci*, 32(25):8649–8662.
- Murphy, K., Bodurka, J., and Bandettini, P. A. (2007). How long to scan? the relationship between fmri temporal signal to noise ratio and necessary scan duration. *Neuroimage*, 34(2):565–574.
- Murray and Bussey (1999). Perceptual-mnemonic functions of the perirhinal cortex. *Trends Cogn Sci*, 3(4):142–151.
- Naya, Y., Yoshida, M., and Miyashita, Y. (2003). Forward processing of long-term associative memory in monkey inferotemporal cortex. *J Neurosci*, 23(7):2861–2871.
- Nobre, A., Allison, T., and McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature*, 372:260–263.
- Nobre, A. and McCarthy, G. (1995). Language-related field potentials in the anterior-medial temporal lobe: II. effects of word type and semantic priming. *J Neurosci*, 15:1090–1098.
- O’Kane, G., Insler, R. Z., and Wagner, A. D. (2005). Conceptual and perceptual novelty effects in human medial temporal cortex. *Hippocampus*, 15(3):326–332.
- Orban, G. A., Zhu, Q., and Vanduffel, W. (2014). The transition in the ventral stream from feature to real-world entity representations. *Front Psychol*, 5:695.

- Parker, A. and Gaffan, D. (1998). Lesions of the primate rhinal cortex cause deficits in flavour-visual associative memory. *Behav Brain Res*, 93(1-2):99–105.
- Patterson, K., Nestor, P. J., and Rogers, T. T. (2007). Where do you know what you know? the representation of semantic knowledge in the human brain. *Nat Rev Neurosci*, 8(12):976–987.
- Rogers, T. and McClelland, J. (2004). *Semantic cognition. A parallel distributed approach*. Cambridge, MA, Massachusetts Institute of Technology Press.
- Saleem, K. S., Suzuki, W., Tanaka, K., and Hashikawa, T. (2000). Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *J Neurosci*, 20(13):5083–5101.
- Seghier, M. L. and Price, C. J. (2011). Explaining left lateralization for words in the ventral occipitotemporal cortex. *J Neurosci*, 31(41):14745–14753.
- Shinkareva, S. V., Malave, V. L., Mason, R. A., Mitchell, T. M., and Just, M. A. (2011). Commonality of neural representations of words and pictures. *Neuroimage*, 54(3):2418–2425.
- Shinkareva, S. V., Mason, R. A., Malave, V. L., Wang, W., Mitchell, T. M., and Just, M. A. (2008). Using fMRI brain activation to identify cognitive states associated with perception of tools and dwellings. *PLoS One*, 3(1):e1394.
- Simanova, I., Hagoort, P., Oostenveld, R., and van Gerven, M. A. J. (2012). Modality-independent decoding of semantic information from the human brain. *Cereb Cortex*.
- Spitsyna, G., Warren, J. E., Scott, S. K., Turkheimer, F. E., and Wise, R. J. S. (2006). Converging language streams in the human temporal lobe. *J Neurosci*, 26(28):7328–7336.
- Suzuki, W. A. and Amaral, D. G. (1994a). Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *J Comp Neurol*, 350(4):497–533.
- Suzuki, W. A. and Amaral, D. G. (1994b). Topographic organization of the reciprocal connections between the monkey entorhinal cortex and the perirhinal and parahippocampal cortices. *J Neurosci*, 14(3 Pt 2):1856–1877.
- Suzuki, W. A. and Naya, Y. (2014). The perirhinal cortex. *Annu Rev Neurosci*, 37:39–53.
- Ueno, T., Saito, S., Rogers, T. T., and Lambon Ralph, M. A. (2011). Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, 72(2):385–396.
- Van Doren, L., Dupont, P., De Grauwe, S., Peeters, R., and Vandenberghe, R. (2010). The amodal system for conscious word and picture identification in the absence of a semantic task. *Neuroimage*, 49(4):3295–3307.
- Vandenberghe, R., Price, C., Wise, R., Josephs, O., and Frackowiak, R. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, 383:254–256.

- Vandenbulcke, M., Peeters, R., Dupont, P., Van Hecke, P., and Vandenberghe, R. (2007). Word reading and posterior temporal dysfunction in amnesic mild cognitive impairment. *Cereb. Cortex*, 17:542–551.
- Vandenbulcke, M., Peeters, R., Fannes, K., and Vandenberghe, R. (2006). Knowledge of visual attributes in the right hemisphere. *Nat. Neurosci.*, 9:964–970.
- Visser, M., Embleton, K. V., Jefferies, E., Parker, G. J., and Ralph, M. A. L. (2010a). The inferior, anterior temporal lobes and semantic memory clarified: novel evidence from distortion-corrected fMRI. *Neuropsychologia*, 48(6):1689–1696.
- Visser, M., Jefferies, E., and Lambon Ralph, M. A. (2010b). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci*, 22(6):1083–1094.
- Visser, M. and Lambon Ralph, M. A. (2011). Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *J Cogn Neurosci*, 23(10):3121–3131.
- Vogel, A. C., Petersen, S. E., and Schlaggar, B. L. (2012). The left occipitotemporal cortex does not show preferential activity for words. *Cereb Cortex*, 22(12):2715–2732.
- Voss, J. L., Hauner, K. K. Y., and Paller, K. A. (2009). Establishing a relationship between activity reduction in human perirhinal cortex and priming. *Hippocampus*, 19(9):773–778.
- Wagner, A., Desmond, J., Demb, J., Glover, G., and Gabrieli, J. (1997). Semantic repetition priming for verbal and pictorial knowledge: a functional MRI study of left inferior prefrontal cortex. *J. Cogn. Neurosci.*, 9:714–726.
- Wang, W.-C., Lazzara, M. M., Ranganath, C., Knight, R. T., and Yonelinas, A. P. (2010). The medial temporal lobe supports conceptual implicit memory. *Neuron*, 68(5):835–842.
- Whitney, C., Kirk, M., O’Sullivan, J., Lambon Ralph, M. A., and Jefferies, E. (2011). The neural organization of semantic control: TMS evidence for a distributed network in left inferior frontal and posterior middle temporal gyrus. *Cereb Cortex*, 21(5):1066–1075.
- Yarkoni, T., Balota, D., and Yap, M. (2008). Moving beyond Coltheart’s n: a new measure of orthographic similarity. *Psychon Bull Rev*, 15(5):971–979.
- Yoshida, M., Naya, Y., and Miyashita, Y. (2003). Anatomical organization of forward fiber projections from area TE to perirhinal neurons representing visual long-term memory in monkeys. *Proc Natl Acad Sci U S A*, 100(7):4257–4262.