Behavioral/Cognitive

Representational Similarity Analysis Reveals Commonalities and Differences in the Semantic Processing of Words and Objects

Barry J. Devereux, Alex Clarke, Andreas Marouchos, and Lorraine K. Tyler
Centre for Speech, Language, and the Brain, Department of Psychology, University of Cambridge, Cambridge, CB2 3EB, United Kingdom

Understanding the meanings of words and objects requires the activation of underlying conceptual representations. Semantic representations are often assumed to be coded such that meaning is evoked regardless of the input modality. However, the extent to which meaning is coded in modality-independent or amodal systems remains controversial. We address this issue in a human fMRI study investigating the neural processing of concepts, presented separately as written words and pictures. Activation maps for each individual word and picture were used as input for searchlight-based multivoxel pattern analyses. Representational similarity analysis was used to identify regions correlating with low-level visual models of the words and objects and the semantic category structure common to both. Common semantic category effects for both modalities were found in a left-lateralized network, including left posterior middle temporal gyrus (LpMTG), left angular gyrus, and left intraparietal sulcus (LIPS), in addition to object- and word-specific semantic processing in ventral temporal cortex and more anterior MTG, respectively. To explore differences in representational content across regions and modalities, we developed novel data-driven analyses, based on k-means clustering of searchlight dissimilarity matrices and seeded correlation analysis. These revealed subtle differences in the representations in semantic-sensitive regions, with representations in LIPS being relatively invariant to stimulus modality and representations in LpMTG being uncorrelated across modality. These results suggest that, although both LpMTG and LIPS are involved in semantic processing, only the functional role of LIPS is the same regardless of the visual input, whereas the functional role of LpMTG differs for words and objects.

Introduction

To what extent are conceptual representations coded in a semantic system common to both words and objects? Although several regions have been identified in multimodal semantic processing, suggesting their relevance to a common semantic system [e.g., left angular gyrus (LAG), fusiform cortex, and lateral and anterior temporal regions; Warrington and Shallice, 1984; Vandenberghe et al., 1996; Bright et al., 2004; Shinkareva et al., 2011], the computational properties of these regions remain poorly understood. This raises the issue of whether regions process the same kinds of semantic representations for both words and objects (indicative of truly amodal semantic processing) or are involved in different kinds of processing depending on stimulus modality. Here we report an fMRI study designed to determine both the commonalties and the differences in the computations subserved for objects and words by different neural regions.

Recently, classifier-based decoding methods have been successfully used to identify regions involved in cross-modal processing by training on data from one modality and testing against data from the other modality (Meyer et al., 2010; Shinkareva et al., 2011; Akama et al., 2012; Simanova et al., 2012; Fairhall and Caramazza, 2013). Although cross-modal classification can reveal invariance across stimulus modality, additional approaches are needed to fully characterize both the similarities and differences across modalities. For example, a region may be involved in semantic processing for both words and objects but as part of different word and object processing networks; therefore, the underlying computations in that region may differ across modality. Furthermore, similar semantic representations may be coded in different regions for each modality. Classification methods are unsuitable for investigating such possibilities, because they depend on spatial and functional isomorphisms between training and testing data.

We use representational similarity analysis (RSA) and searchlight mapping (Kriegeskorte et al., 2006, 2008), which tests how well predicted similarities between stimuli are reflected in the similarities of multivoxel activation patterns to the same stimuli. This allows for a flexible form of pattern analysis, in which the representational dissimilarity matrices (RDMs) for activation within a region may remain the same across modalities, even when the underlying activation profiles of individual voxels differ. Because of this "dissimilarity trick" (Kriegeskorte and Kievit, 2013), RSA is well suited to cross-modal data (even datasets from

different species or imaging modalities; Kriegeskorte et al., 2008). Furthermore, RSA can capture the dissimilarity structure of individual stimuli, as well as distinguish stimulus classes. We use RSA in complementary model- and data-driven analyses, allowing us to characterize the semantic system in several ways. For model-driven RSA, neural dissimilarity patterns are tested against visual and semantic stimulus models, identifying regions in which representational content reflects visual and semantic processing. The data-driven analysis involves clustering searchlight RDMs: RDMs coding for similar representational content will cluster together, regardless of their modality or location. Unlike cross-modal classification, this method can identify representational invariance across regions as well as across modalities. Thus, model- and data-driven RSA permits us to develop a comprehensive picture of both the commonalities and differences in semantic processing for words and objects.

Materials and Methods

Subjects. Fourteen right-handed native British English speakers (four male; median age, 22 years; range, 19–25 years), with normal or corrected-to-normal vision and free of neurological or language disorders, volunteered to take part in the study. All subjects gave informed consent. The study was approved by the Cambridge Psychology Research Ethics Committee.

Stimuli. The study used a total of 60 concepts presented as both pictures and words. Stimuli were 10 concrete object pictures (and corresponding nouns) from each of six common semantic categories: (1) animals (all land mammals); (2) clothing; (3) insects; (4) tools; (5) vegetables; and (6) vehicles. Pictures were 460×460 pixel jpeg images with the object presented against a white background, presented at 7.5° visual angle. To ensure that all of the presented stimuli could be identified reliably at the category level, 10 participants who did not take part in the fMRI study took part in a pretest that included a larger set of items and categories. Pretest participants were instructed to name the category membership of the object ("say the kind of thing that each object is") as in the main fMRI study (see below). We determined the two most common acceptable category responses for items in each category (e.g., "clothes" and "clothing" for the clothing category, "transport" and "vehicle" for the vehicle category). The 60 experimental items were consistently given one of these two most common category names (93.2% of pretest responses). Although the experiment did not require basic-level naming, we excluded pictures that were known to be difficult to identify at the basic level, based on name agreement data from a previous behavioral study (Taylor et al., 2012). The 56 pictures in common with the previous study had high name agreement (average of 85.3%), indicating that stimuli were consistently identified at the basic level.

The items in each of the six categories were also matched on a range of lexical and visual variables. All six categories were matched on the lemma frequency and the number of phonemes in the word [CELEX; (Baayen et al., 1993)], the familiarity of the concept, and the exemplarity of the concept (from behavioral pretests with separate subjects; ANOVAs, all $p > 0.33$). We calculated two measures of visual complexity: (1) the compressed jpeg file size (Székely and Bates, 2000; Bates et al., 2003); and (2) the number of pixels composing the object (Snodgrass and Corwin, 1988; Taylor et al., 2012). Given the inherent visual differences between tools and animals, with animals tending to be more visually complex, we did not attempt to match all categories for visual complexity; instead, we matched pairs of nonliving categories to living categories (following Moss et al., 2005). In particular, for both file size and number of object pixels, animals were matched to vehicles (t tests, both $p > 0.86$), and tools were matched to vegetables (both $p > 0.14$), thus avoiding confounding visual complexity with semantic domain.

Procedure. The stimuli were presented in four scanning sessions, with word stimuli presented in the first two sessions and object stimuli presented in the second two sessions. Words were always presented before objects to reduce the likelihood of eliciting the visual image of the object when seeing the word. Each trial consisted of a centrally presented fixa-tion cross on a white background for 500 ms, followed by the stimulus for 500 ms, followed by a blank screen lasting between 2 and 12 s (timings defined using optseq; http://surfer.nmr.mgh.harvard.edu/optseq/). The mean duration of the intertrial intervals (blank screen) was 3567 ms. Participants were instructed to name the category of the words and objects ("say the kind of thing that each object is"). Each of the four blocks contained three separate runs, with all 60 items presented once in each run and with a blank screen lasting 20 s between runs. Each run lasted 274 s. The order of items was randomized within each run, subject to the constraint that no pair of stimuli occurred consecutively in more than two of the six runs for any participant (to avoid colinearity in the BOLD response for item regressors across runs). Independent pseudorandomizations were created for each participant. The presentation and timing of stimuli was controlled using E-prime 2 (Psychology Software Tools).

fMRI data acquisition. Participants were scanned on a 3 T Tim Trio (Siemens) at the Medical Research Council Cognition and Brain Sciences Unit (Cambridge, UK). fMRI data in the four functional scanning sessions were collected using a gradient-echo echo-planar imaging (EPI) sequence (acquisition time, 2000 ms; TR, 2000 ms; TE, 30 ms; flip angle, 78°; matrix size, 64×64; resolution, 3×3 mm; 32 slices in descending order; 3 mm thick, with 0.75 mm slice gap). Each of the four functional scanning sessions lasted 890 s. We also acquired T1-weighted MPRAGE scans (1 mm isotropic resolution) from each participant.

fMRI preprocessing. Preprocessing was performed with SPM8 (Wellcome Institute of Cognitive Neurology, London, UK; www.fil.ion.ucl.ac.uk). Preprocessing involved slice-time correction and spatial realignment of all EPI images in each session to the first image in the first session (excluding four initial lead-in images, which were removed from each session). Images were not spatially normalized or smoothed, in order to take advantage of high-spatial-frequency pattern information within each participants' data in the RSAs (Kriegeskorte et al., 2006). The preprocessed images for each participant were analyzed using the general linear model (GLM), with words and pictures analyzed separately. Both analyses used separate regressors for each concept (i.e., 60 stimulus regressors). Also included were six head motion regressors for each session, a regressor for the difference in the global mean for the two sessions, and 13 regressors for each session, based on the basis functions of a discrete cosine transform, to capture low-frequency trends (minimum frequency, $\frac{1}{128}$ Hz). To ensure maximal coverage of the anterior temporal lobes, the SPM implicit masking threshold was set to 10% of the mean signal over all voxels in the brain (compared with a default SPM masking threshold of 80%). From this GLM analysis, we obtained a single β-image for each word and picture. To create a gray-matter mask for each subject, the structural image was coregistered to the functional data and was segmented and normalized to Montreal Neurological Institute (MNI) space using unified segmentation normalization as implemented in SPM8. For each subject, the native space gray-matter probabilistic map was resliced to the functional data and thresholded at one-third to create a binary gray-matter mask for use in RSA searchlight mapping (see below).

RSAs. RSA was used to identify the representational content of multivariate activation patterns across the cortex. RSA involves computing a second-order correlation (typically Spearman's correlation) between model RDMs and activation-pattern RDMs (Kriegeskorte et al., 2006, 2008; Mur et al., 2009). Model RDMs exemplify the similarity between items as predicted by a computational model or some hypothesis about how the stimulus space is structured, whereas activation-pattern RDMs are computed for a (typically spatially contiguous) set of voxels using some dissimilarity function (typically $1 -$ Pearson's correlation across voxels). In this way, RSA permits testing of predictions against the data in a manner that abstracts away from the underlying representational substrate (i.e., stimulus attributes in the case of a model and voxel values in the case of fMRI data).

We constructed several model RDMs based on visual and semantic properties of the 60 stimuli. For low-level visual information, we used RDMs based on image silhouettes, which are known to be highly predictive of activation patterns in early visual cortex (Kriegeskorte et al., 2008). Image bitmaps were binarized (pixels in object $= 1$; white background pixels $= 0$), and distances were computed between them using 1-Jaccard

similarity (Fig. 1A). To create silhouette models for the words in the same way as for the pictures, we constructed a continuous edge outlining each word, preserving ascender and descender information (Fig. 1B). Silhouette distance between words was calculated with respect to these edge outlines as 1-Jaccard similarity, in the same way as computed for objects. The semantic category structure was coded by an RDM in which pairs of items from the same category were similar and pairs of items from different categories were dissimilar (Fig. 1C). These model RDMs were compared with activation RDMs using a standard multivariate searchlight technique (spherical searchlights; radius, 10 mm) across all gray-matter voxels for each participant, producing a Spearman's correlation map for each subject and model. The RSA was performed using the MRC-CBU RSA toolbox (revision 103) for MATLAB (http://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes) and custom MATLAB scripts. For group random-effects analyses, the Spearman's correlation maps for each participant were Fisher transformed, normalized to standard MNI space, and spatially smoothed with a 10 mm FWHM Gaussian kernel. To ensure optimal control of type I error, group-level random-effects analyses were conducted by permutation testing with the SnPM toolbox (Nichols and Holmes, 2002; http://go.warwick.ac.uk/tenichols/snpm). Variance smoothing of 10 mm FWHM and 10,000 permutations were used in these analyses. We report voxel-level familywise error (FWE) corrected p values < 0.05 (one-sample pseudo-t statistic). For visualization, the volumetric statistically thresholded maps were mapped onto the PALS-B12 surface atlas in CARET version 5.6.

Standard RSA allows us to compare similarity of activation patterns to similarity patterns predicted by a computational model of the stimuli. Because we are interested here in the commonalities and differences in word and object processing, we examine the relationship between word and picture representations directly as well as through the prism of their respective relationships to predicted model structures. Indeed, comparing correlations with model RDMs for words and pictures is not sufficient to determine commonality in representations for words and pictures: activation RDMs for words and pictures may both be correlated with a given model RDM yet may not be strongly correlated with each other. Therefore, following the standard RSA with visual and semantic model RDMs, we also conducted a data-driven version of RSA in which activation RDMs drawn from one modality were tested against the independent data in the other modality using searchlight mapping. This allows us to investigate whether representational content from a given region in one modality predicts the representational content of the same region (or, indeed, a different region) in the other modality.

Representational cluster analyses. We also developed a custom analysis, which we call representational cluster analysis (RCA), which allows us to determine and visualize the representational similarities across the entire set of activation patterns for the two presentation modalities (Fig. 2). Intuitively, the goal of this cluster analysis is to partition the searchlight spheres centered at every voxel from the word and picture data into groups with similar activation RDMs. RDMs reflecting a particular kind of representational content should all be similar and therefore cluster together. For example, voxels in the picture data that are sensitive to object shape should have similar RDMs and thus cluster together. Connolly et al. (2012) used a similar approach to cluster RDMs within a single modality, with the goal of creating functional ROIs for subsequent analyses. Of particular interest to us are cases in which clusters contain voxels from both modalities, reflecting representational content that occurs for both words and pictures. Critically, this analysis method is blind with respect to whether the voxels showing common representational content are spatially corresponding; the question of whether shared representational content is found in corresponding regions in the two modalities or in different regions is determined from the data via the clustering. This can be contrasted with cross-modal classifier-based methods, which pre-
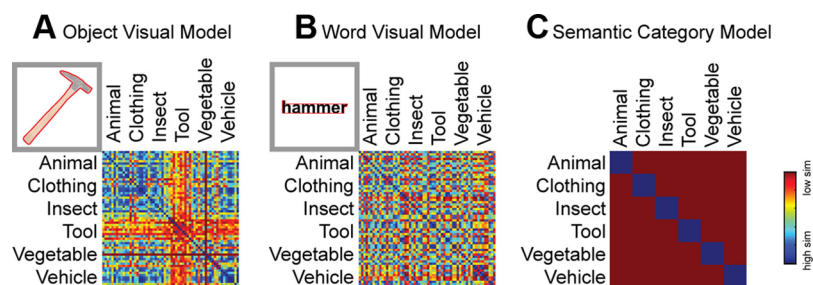
Figure 1. The three model RDMs used for RSA. *A*, Object visual silhouette model, in which distances are calculated with respect to overlap of object silhouettes, such as that depicted for the example of hammer (red line depicts silhouette outline). *B*, Word visual silhouette model, with distances calculated with respect to word outlines. *C*, Semantic category model.

suppose that common representational content is only found in corresponding voxels across modalities.

To ensure maximal overlap in RDMs across subjects, activation RDMs were calculated over the whole brain (rather than within the subject-specific gray-matter masks used in the standard RSA). We combined activation RDMs across subjects by mapping each subject's activation RDMs into the standard MNI space. RDMs at each voxel coordinate in MNI space were then averaged across subjects (excluding any voxels for which data from one or more participants was missing), for the word and picture data separately, and a cortical mask [based on the Automated Anatomical Labeling Atlas (Tzourio-Mazoyer et al., 2002)] was applied. The resulting 72,483 RDMs (36,236 for words and 36,247 for pictures) were reshaped to vectors. To reduce redundancy in the 1770-dimensional vectors and to speed clustering, these vectors were reduced to 200 dimensions through singular value decomposition (SVD). The SVD transformation preserves almost all information about distances in the original high-dimensional space (Pearson's correlation between Euclidean distances in original space and Euclidean distances in 200-dimensional space, $r = 0.98$). The reduced vectors were entered into a k-means cluster analysis, undertaken with the SciKit-Learn package (Pedregosa et al., 2011) in Python. Euclidean distance was used as the second-order comparison metric in clustering (as required by the k-means algorithm), and the best clustering solution from 200 runs with different cluster centroid seeds was saved for each cluster analysis. Separate cluster analyses were performed for every number of clusters from 2 to 28, and the optimal number of clusters was selected on the basis of the proportion of variance explained by each cluster solution (see Results). Finally, maps depicting the membership of RDMs in clusters were created on the basis of the MNI voxel coordinates of the RDMs.

Results

RSA results

Participants were very accurate at providing appropriate category labels for each stimulus (97.9% of responses), indicating that they understood the task and processed the stimuli semantically. In the RSA searchlight mapping, visual silhouette and semantic model RDMs were tested against the word and picture data separately. Figure 3 presents the results for the picture data. The visual silhouette RDM showed significant correlations with early visual cortex, consistent with previous results using similar model RDMs in RSA (Kriegeskorte et al., 2008). The semantic category RDM correlated significantly with activation patterns extensively throughout left temporal cortex, with peak effects in the left posterior middle temporal gyrus (LpMTG), extending into lateral occipital cortex (LOC), the inferior parietal lobule (LAG and the posterior supramarginal gyrus), and the left intraparietal sulcus (LIPS), as well as left posterior superior temporal gyrus (STG), lateral and medial fusiform, lingual gyrus, and medial occipital cortex (Table 1). More anteriorly, the cluster extended into the precentral and postcentral gyrii, middle frontal gyrus (MFG), and left inferior frontal gyrus (LIFG; BA 44),
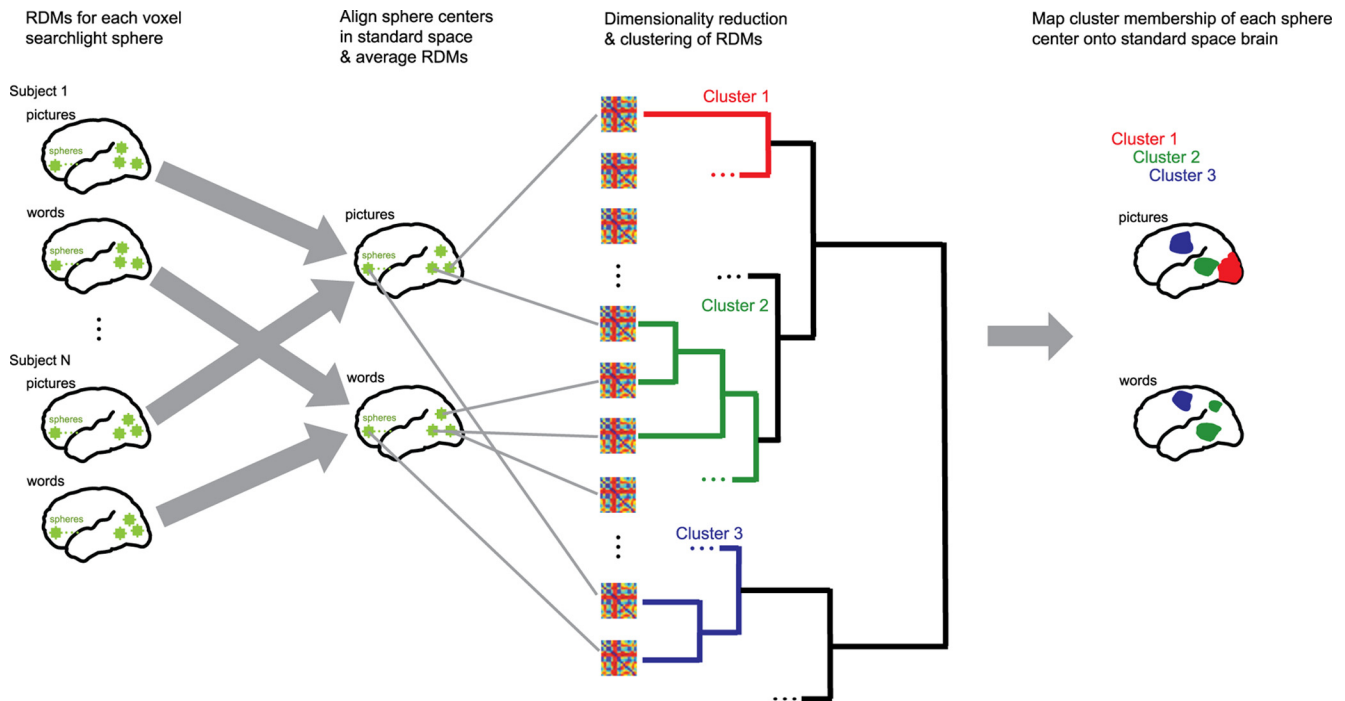
Figure 2. Analyzing commonalities and differences of word and picture processing by clustering of RDMs. RDMs for each voxel searchlight sphere are aligned in standard space and averaged across participants. The resulting 72,483 RDMs are reduced to 200-dimensional vectors by SVD. These vectors in 200-dimensional space are repeatedly clustered by k-means clustering to obtain an optimal number of clusters (see Materials and Methods). (Note that clustering is visually illustrated with a dendrogram here, although k-means clustering is not hierarchical). RDMs which are similar, and thus reflect commonality in the underlying multivoxel patterns, will tend to cluster together. Some kinds of patterns may be specific to a single modality, with all cluster members being RDMs drawn from one or the other modality. For example, RDMs from early visual cortex for the picture data may all reflect low-level visual information in the object images and thus cluster together (Cluster 1, red), but it is unlikely that any RDMs from the word data will reflect low-level visual information in the picture silhouettes. Other clusters may contain RDMs from both modalities, reflecting pattern information that is invariant with respect to the mode of presentation (e.g., clusters 2 and 3). Clusters spanning modalities may include RDMs from corresponding cortical regions for the two modalities (e.g., Cluster 3, blue), reflecting activation patterns that are unaffected by the mode of presentation (e.g., activation associated with the verbal response, which was the same for words and pictures). Other clusters spanning modality may include RDMs from non-overlapping cortical regions (e.g., Cluster 2, green), reflecting similar representational content being processed in different areas across modalities.

whereas a smaller cluster also included left middle frontal cortex, extending into BA 45. Right-hemisphere effects were weaker and lesser in extent than those found on the left. The strongest correlation was found in a cluster spanning the precentral, postcentral, and supramarginal gyrii. Contralateral to the strongest effects in the left hemisphere, there were significant correlations in right fusiform, right posterior MTG, and right precuneus. Significant correlations were also found in posterior cingulate and superior medial frontal cortex.

Figure 4 presents the results for the word data in which the visual silhouette model correlated significantly with a single cluster in the left occipital pole. The semantic category RDM correlated significantly with activation patterns extensively throughout left temporal cortex, with a peak effect in the left inferior parietal lobule (Table 2). This large cluster included most of the AG and the LIPS and extended into left posterior supermarginal gyrus, left STG (LSTG), and LMTG. This cluster also extended anteriorly over portions of the left postcentral gyrus, precentral gyrus, and MFG and into LIFG (BA 45). In the right hemisphere, activation was also found in postcentral gyrus, precentral gyrus, and MFG, BA 45, and in small areas of the superior frontal gyrus and the inferior parietal lobule. Significant correlation was also found in the precuneus and posterior cingulate.

The RSAs revealed extensive left-hemisphere effects of semantic category for both modalities (Fig. 5). The clearest difference in the two sets of results is the absence of word effects in the fusiform and most of LOC. Effects for words extend more anteriorly into MTG and STG and show more widespread effects in left frontal

cortex. Overall, however, the pattern is one of considerable overlap for the semantic model for words and pictures, including LpMTG, left anterior LOC, LAG, and LIPS.

To characterize the relationships between semantic representations in different regions, both within and across modality, we next examined the representational content in the areas corresponding to the peak semantic category RSA effects for words and pictures, which were found in the LIPL (the lateral bank of the LIPS) and LpMTG, respectively (Tables 1, 2). In total, we extracted four average activation RDMs, from the two peak locations in both the word and picture data, and examined the correlations between RDMs across modality (Fig. 6B). For the LIPS peak, there was a significant correlation between the word and picture RDMs (Spearman's $\rho = 0.207$, $p = 0.002$; p values from a permutation test with 10,000 permutations of RDM row/column labels, Bonferroni-corrected for the six pairwise RDM comparisons), indicating invariance of the representational profile of this voxel across words and pictures. However, for the LpMTG peak, the word and picture RDMs were not significantly correlated (Spearman's $\rho = 0.078$, $p = 0.540$; Bonferroni-corrected), although LpMTG shows sensitivity to semantic category information for both words and pictures (Fig. 5). Interestingly, there is a significant cross-modal, cross-region correlation: patterns in LpMTG for pictures are similar to patterns in LIPS for words (Spearman's $\rho = 0.216$, $p = 0.003$; Bonferroni-corrected).
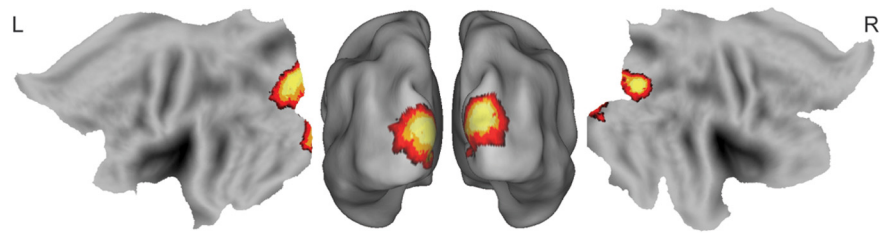
We next used the peak activation RDMs to conduct a cross-modal seed voxel correlation analysis (Fig. 6A,C) that aims to

identify regions with modality-invariant semantic processing. The goal of this analysis is to examine whether the representational content of a region maximally sensitive to semantic category in one modality predicts the representational content of the same region in the other modality. For this analysis, both the semantic peak RDM from the picture data, located in LpMTG, and the semantic peak RDM from the word data, located in LIPL/LIPS, were treated as predictor RDMs and correlated with the activation data from the other independent dataset, using the same standard searchlight mapping RSA method that we used with the semantic category and visual silhouette models. This analysis can be regarded as a multivoxel pattern analysis (MVPA) searchlight mapping analog of β-series correlation (Rissman et al., 2004), in which the set of item pair correlations calculated within a searchlight sphere replaces the univariate β-series of a voxel. The results of this analysis corroborated the peak voxel analyses described above, revealing differences in word and picture semantic processing. The LpMTG RDM from the picture data yielded a significant cluster in the analysis on the word dataset, centered on the LIPL with peak located in superior LAG, extending into superior anterior occipital cortex, the LIPS, and marginally into the left superior parietal lobule (LSPL) (Fig. 6A). Two smaller clusters were located in left MFG. Interestingly, there were no effects in the LpMTG region in the word data, although the seed RDM comes from this region in the picture data. We used the same approach for the analysis of the picture data, using the LIPS peak semantic RDM from the word data as the predictor RDM (Fig. 6C). The peak effect was in LpMTG, in a cluster extending into inferior left LOC and posterior left inferior temporal gyrus. Notably, this cluster did not extend to many of the regions showing significant semantic category effects for pictures, including much of the AG, the fusiform, and more anteriorly along the MTG. A second cluster is centered on left superior occipital cortex and the LIPS, extending marginally into the adjacent cortex in LIPL and LSPL. This cluster has a high degree of overlap with the dorsal areas showing semantic category effects for both words and pictures (Fig. 5) and the significant effect of the picture RDM seed in the word data (Fig. 6A), indicating that the LIPS contains representations that are relatively consistent across the two stimulus modalities. Furthermore, the LIPS seed did not show significant effects in ventral motor cortex, suggesting that the invariance of LIPS cannot be accounted for by the common verbal response for words and pictures.
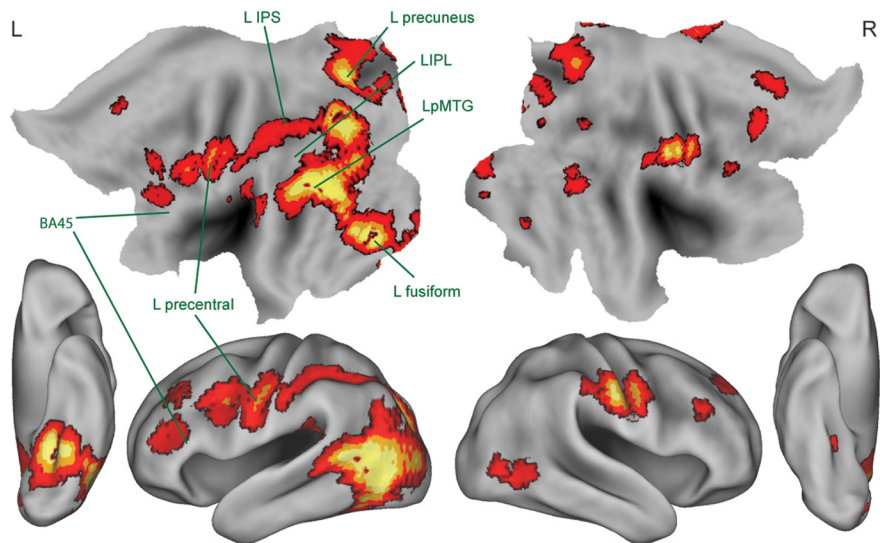
RCA results

RCA partitions activation RDMs, across both the picture and word data, into clusters that show similar patterns of representational content. Of particular interest are clusters spanning the two datasets, which indicate common representations found in both modalities. Areas belonging to such cross-modal clusters

Figure 3. Significant effects in RSA on picture data ($p_{FWE} < 0.05$). **A**, Results for the visual silhouette model. **B**, Results for the semantic category model. L, Left; R, right.
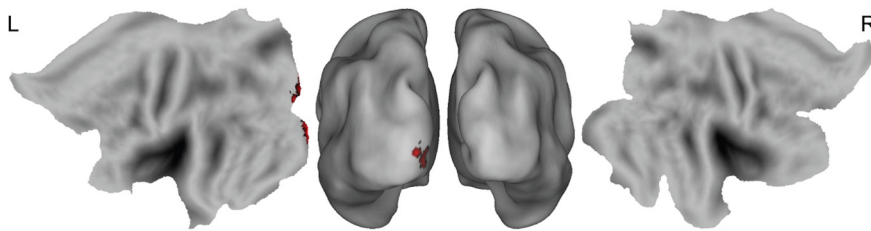
Table 1. MNI coordinates, significance levels, and anatomical regions for peak RSA searchlight effects for the semantic category model for pictures

| Region	Cluster extent	p_{FWE}-corr	Pseudo-t	x	y	z
Left hemisphere						
LpMTG	3143	0.0001	8.47	−48	−64	6
Left fusiform		0.0004	7.91	−27	−49	−9
Left superior middle occipital		0.0007	7.35	−27	−73	29
Left inferior middle occipital		0.0015	6.77	−36	−88	6
Left postcentral		0.0053	5.95	−51	−10	36
Left calcarine		0.0086	5.66	−6	−58	10
LIPL		0.0126	5.43	−36	−43	44
Left middle frontal	97	0.022	5.07	−36	32	32
LIFG (BA45)		0.0274	4.89	−45	38	10
Right hemisphere						
Right postcentral	258	0.0012	6.89	60	−16	36
Right posterior cingulate	73	0.0058	5.91	12	−28	44
Right superior frontal	50	0.0077	5.75	21	44	40
Right inferior temporal	92	0.01	5.58	48	−61	−5
Right superior medial frontal	146	0.0116	5.5	0	35	40
Right calcarine	73	0.0128	5.41	12	−85	6

Effects in clusters with extent <20 voxels not shown.

need not be the same as those regions showing significant correlations with the semantic model for both words and pictures identified above because (1) word and picture RDMs may be correlated with each other even if they are not correlated with the semantic model, and (2) word and picture RDMs may be correlated with the semantic model but not with each other. Furthermore, unlike classification-based approaches that rely on training

A Visual Silhouette RDM

B Semantic category RDM
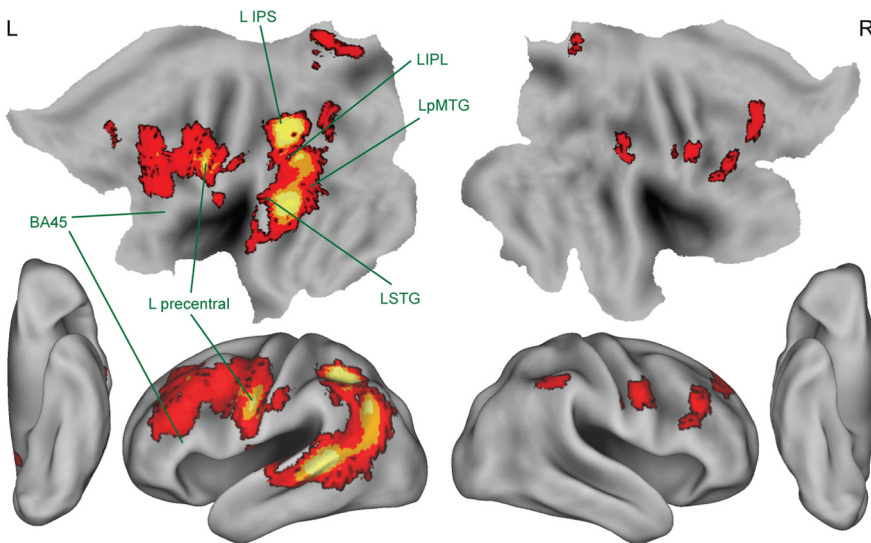
Figure 4. Significant effects in RSA on word data. **A**, Results for the visual silhouette model. **B**, Results for the semantic category model. L, Left; R, right.

Table 2. MNI coordinates, significance levels, and anatomical regions for peak RSA searchlight effects for the semantic category model for words

Region	Cluster extent	p_{FWE}-corr	Pseudo-t	x	y	z
Left hemisphere						
LIPL	1957	0.0001	7.86	−36	−52	48
LMTG		0.0008	7.25	−57	−43	2
Left precentral		0.001	6.68	−54	−7	32
LpMTG		0.0041	5.87	−48	−67	18
Left middle frontal		0.0055	5.67	−36	23	44
LIFG (BA 45)		0.0113	5.29	−45	29	21
Left precentral		0.0117	5.27	−39	−4	48
L posterior cingulate	127	0.0145	5.11	−6	−52	32
Left calcarine		0.0329	4.59	−6	−61	14
Right hemisphere						
Right IFG (BA45)	61	0.0184	4.97	51	26	21
Right precentral	72	0.0191	4.95	54	2	36
Right superior frontal	55	0.0202	4.9	21	44	40
RIPL	32	0.0295	4.67	51	−46	48

Effects in clusters with extent <20 voxels not shown. RIPL, Right inferior parietal lobule.

and testing on spatially corresponding voxels across modalities (Shinkareva et al., 2011; Fairhall and Caramazza, 2013), clustering allows us to identify and visualize more general forms of correspondence across modalities, such as cases in which different areas in the two modalities code for the same representational content.

k-Means clustering solutions were generated for all values of *k* (i.e., number of clusters) between 2 and 28, and the proportion variance explained by each solution was recorded. The *k* = 10

cluster solution was selected for additional analysis on the basis of the scree test (Cattell, 1966) and because 10 clusters explain 80% of the variance in the RDM space (however, note that similar clustering results are found with similar numbers of clusters).

Three of the 10 clusters span both word and picture data and show interesting patterns of similarity and difference in the representational topology for the two modalities (Fig. 7). Cluster 1 includes almost identical areas of cortex for both modalities, centered bilaterally on ventral sensorimotor cortex associated with speech-articulation representations (Guenther et al., 2006; Bouchard et al., 2013). Because participants' verbal category responses for each concept tended to be the same regardless of whether the concept was presented as word or picture, the consistency of this cluster across modalities likely reflects the common similarity structure associated with articulatory feature information corresponding to these verbal responses. Cluster 2 also includes areas of bilateral sensorimotor cortex for both words and pictures, in voxels surrounding cluster 1 (because the *k*-means algorithm uses Euclidean distance as the RDM comparison metric, clusters must be similar in terms of the magnitude of RDM values as well as their pattern similarity; for this reason, clusters tend to form bands of cortex reflecting similar overall magnitude as well as similar multivariate patterns). However, in other areas, membership in cluster 2 differs for words and pictures. For words, cluster 2 is found in left supramarginal gyrus, LIPS, and the supplementary motor area, bilaterally. For the picture data, cluster 2 is found in the fusiform, bilaterally, as well as LOC bilaterally, but not in left supramarginal gyrus or LIPS. In other words, there are voxels in bilateral fusiform and LOC for pictures for which the dissimilarity structure is closer to the dissimilarity structure of left supramarginal and LIPS for words than to the dissimilarity structure found in the fusiform and LOC for words. Cluster 3 is a larger cluster that again includes many of the regions that were significant in the RSAs with the semantic category model. In particular, cluster 3 includes voxels from the left and right fusiform, occipital cortex, superior and inferior parietal lobules, and bilateral sensorimotor cortex, in both modalities.

Although the LpMTG and LAG showed significant correlations with the semantic category RDM for both modalities (Fig. 5), the cluster analysis reveals differences in the activation RDMs of these regions for words and pictures. Voxels in LpMTG and LAG are not members of any of the three cross-modal clusters, despite LpMTG and LAG showing strong semantic category sensitivity in both modalities. This is consistent with the cross-modal comparison of peak semantic effects (Fig. 6), which revealed no significant correlation between the LpMTG RDM for words and the LpMTG RDM for pictures. The cluster analysis also supports our seed-based data-driven RSA, in which the peak LIPS RDM for words showed significant effects in LpMTG and inferior anterior left LOC for pictures. This pattern is similar to the pattern

of cross-modal cluster membership observed for cluster 2. Furthermore, the cluster analysis shows that LIPS belongs to cross-modal clusters, but LpMTG does not. Finally, the region of cluster 2 membership in LIPS in the word data has a high degree of overlap with the dorsal areas showing semantic category effects for both words and pictures (Fig. 5) and the significant effect of the word RDM seed in the picture data (Fig. 6C). Together, these findings again indicate that the LIPS processes representations that are relatively invariant to stimulus modality. Therefore, the cluster analysis reveals subtle differences in the representational topology of regions involved in the semantic processing of words and pictures, with LOC voxels for pictures clustering with LIPS voxels for words and the absence of cross-modal clustering in LpMTG and LAG. This latter result is particularly surprising given that these regions had high sensitivity to semantic category for both modalities.

Finally, we checked whether the cross-modal differences revealed by the RCA were also present when RDMs are compared with metrics other than Euclidean distance (used in the k-means clustering algorithm). We constructed the full $72,483 \times 72,483$ second-order dissimilarity matrix of all the RDMs that had been entered into the cluster analysis (using $1 -$ Spearman's correlation as the second-order dissimilarity measure, as is conventional in RSA). We then interrogated this matrix to find the cross-modal pair of voxels that had the lowest dissimilarity of all voxels drawn from the two modalities. These maximally similar points were in areas very highly sensitive to semantic category, at MNI coordinate $(-30, -52, 40)$ in the word data (in the LIPS just medial to the peak semantic category effect for words) and $(-45, -67, -5)$ in the picture data (in anterior inferior left LOC). This again demonstrates commonality of representational structure for LIPS for words and the left temporo-occipital region for pictures, corroborating our interpretation of cluster 2 in the RCA analysis and demonstrating that the cross-modal differences revealed by the RCA are present regardless of whether RDMs are compared with Euclidean distance (as used in the k-means clustering algorithm) or Spearman's correlation (as conventionally used in RSA).

Discussion

In this study, we used complementary model-based and model-free RSA to determine the extent to which a common semantic system is evoked by words and objects. This entailed both identifying regions sensitive to semantic category information in the two modalities and elucidating potential differences between them. The semantic category model showed extensive left-lateralized effects in both modalities, with considerable overlap, most notably in LpMTG, LAG, and LIPS. Model-free RSA using word and picture seed RDMs taken from the LpMTG and LIPS showed that the LpMTG RDM from the object data correlated with the LIPS in the word data rather than the LpMTG in the word data, indicating substantial differences in LpMTG representational content across modalities. The seed-based analysis was extended to the whole brain with RCA, which describes the representational topography across the brain for both datasets. This
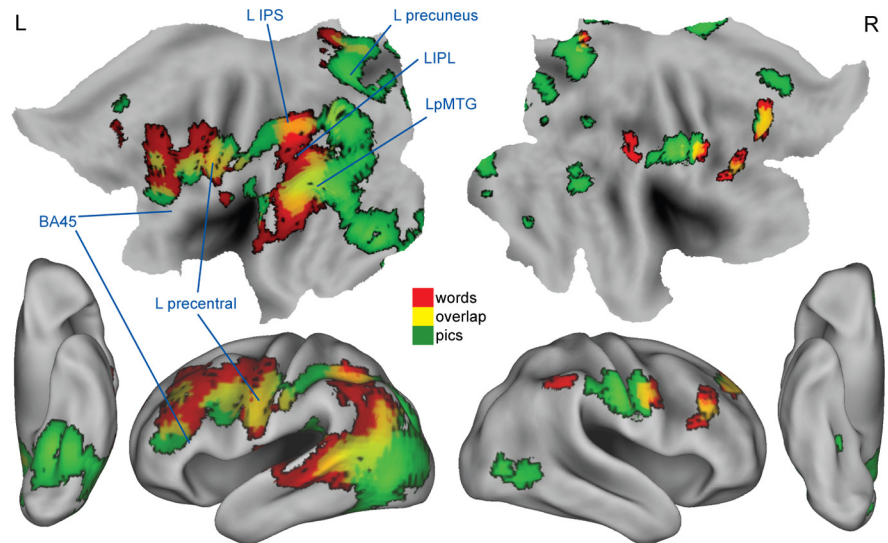

Figure 5. Significant RSA results with the semantic RDM for words, pictures, and their overlap. L, Left; R, right.

revealed complex relationships between regions across modalities. First, we found cross-modal, cross-regional clustering, with representations in LIPS for words and LOC for objects clustering together, consistent with the seed-based RSA analyses. Furthermore, LpMTG was not included in the three clusters that spanned modalities, indicating that representations computed in LpMTG were relatively modality specific, again consistent with the seed-based analyses and indicative of differences in LpMTG across modalities. Together, these results suggest that LpMTG, LAG, and LIPS are sensitive to semantic processing for both words and objects; however, although representations in LIPS are relatively invariant across modality, representations in LpMTG differ, a key finding made possible by combining model- and data-driven RSA approaches.

Previous characterizations of the semantic system have typically focused on identifying regions involved in amodal semantics rather than on the specific computations performed by different regions. The multi-modal regions we identified include many regions known to be engaged in semantic processing for words and objects separately (for overviews, see Martin, 2007; Binder et al., 2009). Moreover, the overlapping semantic category effects in LpMTG, LAG, LIPS, and the posterior cingulate are broadly consistent with univariate and classifier-based MVPA decoding methods targeting multi-modal invariance (Chee et al., 2000; Shinkareva et al., 2011; Visser et al., 2012; Fairhall and Caramazza, 2013). However, the lack of second-order correlation and clustering of representations for words and objects in LpMTG, despite the involvement of this region in semantic processing for both modalities, supports a more differentiated account in which the functional role of LpMTG in the two modalities is quite different, with spatially overlapping but distinct processing within LpMTG giving rise to different multivoxel patterns for words and objects.

Our LpMTG effects for semantic category in the object data are in an area known to be selective to manipulable manmade objects, in particular tools (Mahon et al., 2007), and LpMTG may be critical to understanding specific aspects of the meaning of an object, such as motion information associated with the object (Chao et al., 1999; Beauchamp et al., 2003). The LpMTG and LAG have strong functional connectivity with ventral temporal cortex and may function as part of a distributed semantic network that integrates visual form characteristics of viewed objects
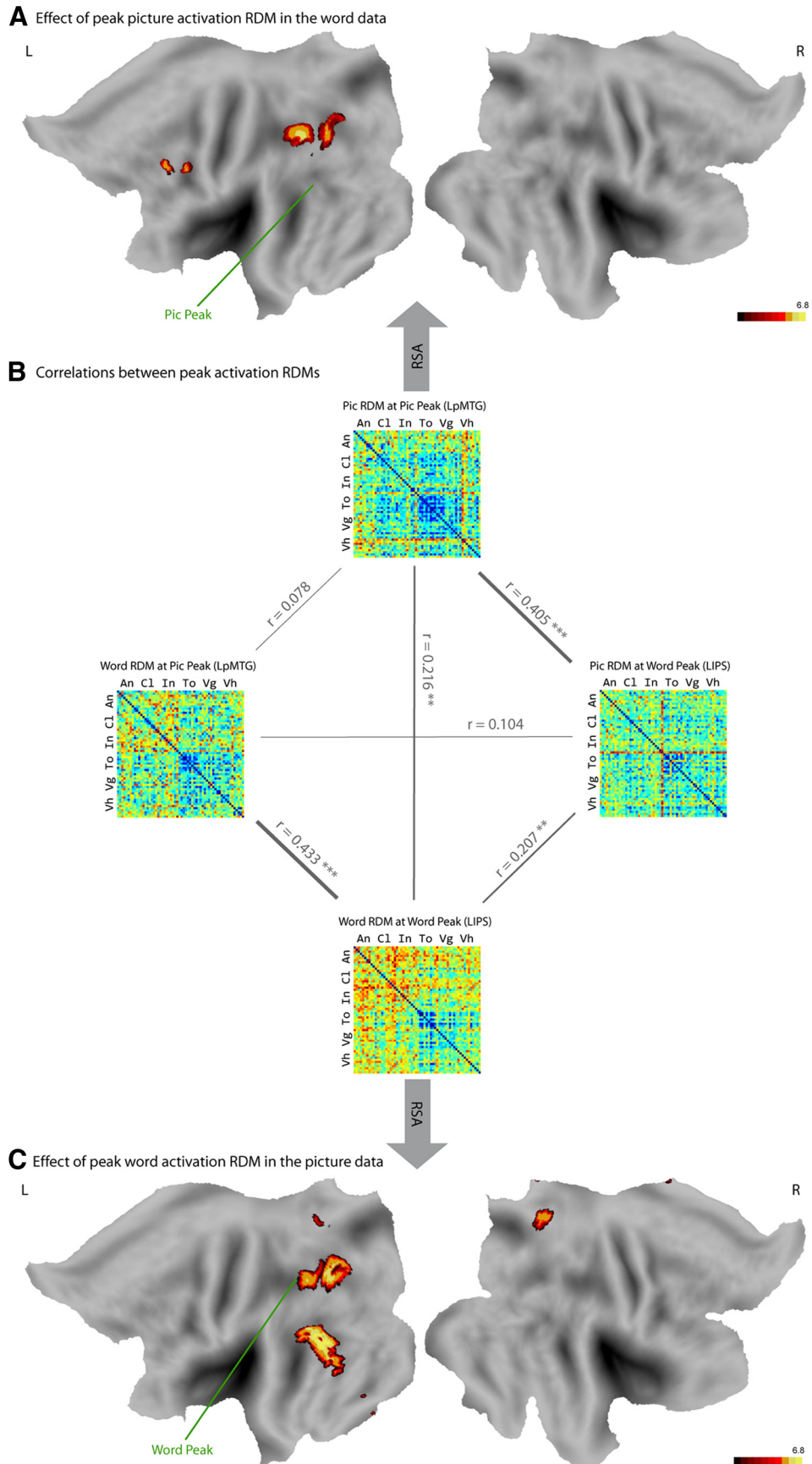
Figure 6. Cross-modal seed voxel correlation analysis. The RDMs from each subject at the location of the peak semantic effects for words and pictures are extracted and averaged, and the four resulting RDMs are cross-correlated [Spearman's ρ (r); **B**]. The 10 rows and columns corresponding to each of the six semantic categories are labeled (An, animal; (*Figure legend continues.*)

with information about object motion and use (Mahon et al., 2007). Thus, multivoxel patterns in LpMTG for objects may be dependent on visual object processing in the ventral stream, which is absent for words. For words, the role of the LpMTG in semantics may relate to separate processes of lexical access, such as mapping between word forms and meaning representations (Badre et al., 2005; Lau et al., 2008). A wealth of neuroimaging evidence implicates LpMTG and LAG in semantic retrieval of lexical representations as part of a left-lateralized language network, for both spoken and written words (Dronkers et al., 2004; Vigneau et al., 2006; Humphries et al., 2007; Tyler and Marslen-Wilson, 2008; Binder et al., 2009; Tyler et al., 2013). On this account, LpMTG is differentially engaged in processes of object motion understanding and lexical–semantic activation for objects and words, respectively, as part of separable object processing and lexical access networks. Critically, this predicts LpMTG involvement in semantic processing for both modalities but with the multivoxel patterns differing across modalities, as observed in our study. This finding has important consequences for research investigating modality-invariant semantics, because it demonstrates that identifying regions involved in both word and picture processing is insufficient to claim that they form part of a common amodal semantic network.

Complementary to representational differentiation for LpMTG, we found relatively invariant representations in a region centered on the lateral bank of the LIPS, with significant semantic category correlation for both words and pictures (Fig. 5), cross-modal correlation of word and picture RDMs (Fig. 6B), and cross-modal cluster membership (Fig. 7). The LIPS is part of the dorsal visual processing stream, with strong connectivity to early visual cortex and sensitivity to visuospatial information (Felleman and Van Essen, 1991; Rizzolatti and Matelli, 2003). However, recent neuroimaging evidence has elaborated on the functional role of the IPS, implicating it in semantic processing that is independent of perceptual input. In particular, the IPS has been implicated in tasks that are semantically demanding or require access to particular feature or category information (Cristescu et al., 2006; Noonan et al., 2013). In a meta-
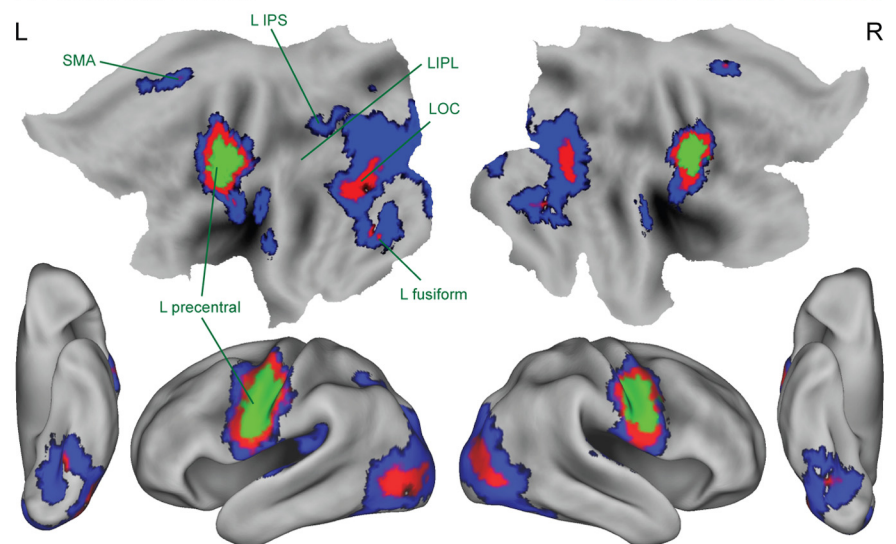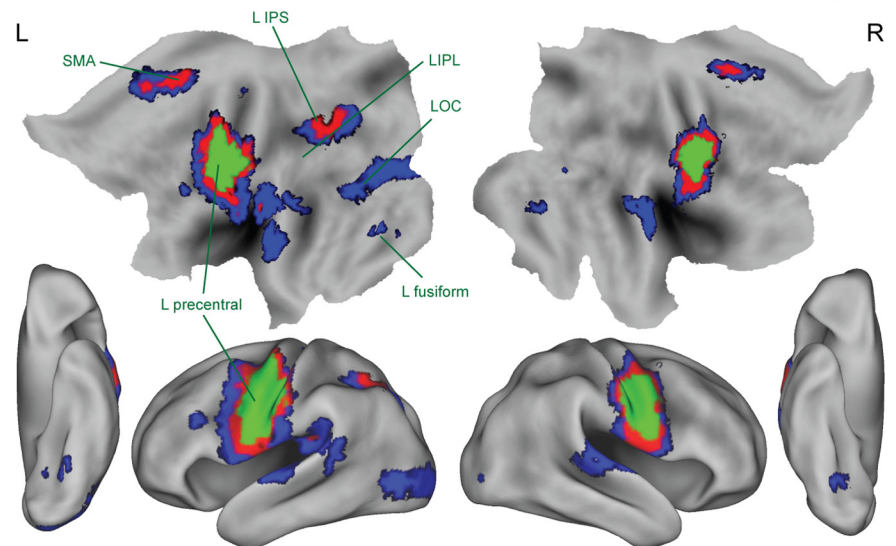
Figure 7. Results of the RCA. The three clusters that cluster together RDMs from both modalities are shown. Non-colored regions belong to clusters specific to a single modality (for details, see Results). L, Left; R, right; SMA, supplementary motor area.

(*Figure legend continued.*) Cl, clothing; In, insect; To, tool; Vg, vegetable; Vh, vehicle). *p < 0.05, **p < 0.01, and ***p < 0.001, significance levels estimated from a permutation test with 10,000 permutations of the 60 RDM row/column labels and Bonferroni-corrected for six comparisons. *A*, Results of RSA on the word data, using the semantic peak RDM from the picture data located in LpMTG [Pic RDM at Pic Peak (LpMTG)], as the predictor RDM. *B*, Results of RSA on the picture data, using the semantic peak RDM from the word data located in LIPL/LIPS [Word RDM at Word Peak (LIPS)], as the predictor RDM. RSA statistical procedures and thresholding as in Figures 3 and 4. L, Left; R, right.

analysis of 53 fMRI studies, primarily using written words as stimuli, Noonan et al. (2013) identified a dorsal angular gyrus (dAG)/IPS region as a key part of network underpinning semantic control processes, responsible for the retrieval of specific semantic information given a particular task or context. This region is highly overlapping with our LIPS region, with its peak lying just anterior to the cross-modal LIPS voxels identified by RCA (Fig. 7). On the assumption that the IPS plays a key role in targeted feature retrieval and that the category judgment task requires access to particular kinds of feature information (i.e., taxonomic information or other semantic features diagnostic of category), then the representational invariance of LIPS suggests that the semantic feature information required for making category judgments does not differ as a function of stimulus modality.

What kinds of semantic feature information may be relevant to semantic category judgments? Related to its function in goal-directed action understanding, the dAG/LIPS has often been identified as an important region in the semantic (as opposed to visual or motor-related) representation of tools, in particular in-

formation about their function and use (Creem-Regehr and Lee, 2005; Creem-Regehr et al., 2007; Mahon et al., 2007, 2010; Valyear et al., 2007; Hoeren et al., 2013). Therefore, one interpretation of our results is that the retrieval of functional information via the LIPS is important for identifying an item as a tool, regardless of whether the input is a word or a picture. This view is consistent with research with congenitally blind people suggesting that the semantic representation of tools in LIPL and LIPS does not explicitly depend on processing visual information (Mahon et al., 2010). On this account, information processed by LIPS is feature-type specific (i.e., functional information) but stimulus-modality independent. However, the extent to which LIPS is involved in general semantic control processes or access of specific functional information about tools must remain a question for future research. The novel contribution of our results is to show that similar representations are computed in LIPS for both words and objects.

Could LIPS representational invariance be explained by the common verbal response across modalities? Although possible, we feel that this interpretation of the results is unlikely. First, LIPS is not typically implicated in speech production. Second, the common verbal response works against one of our key findings: lexical activation of category nouns could be hypothesized to involve LMTG, but we actually find that representations in LMTG are uncorrelated across modality.

Notably absent in our data are semantic effects in anterior temporal cortex (ATC). Although theories differ, ATC is often associated with the processing of amodal or heteromodal semantic representations that are abstracted away from stimulus modality (Taylor et al., 2006; Bright et al., 2007; Patterson et al., 2007; Binney et al., 2010; Holland and Lambon Ralph, 2010; Visser et al., 2012). However, a critical factor in our experiment is that the task required access to category-level rather than item-specific information. Various studies have shown that ATC is critically involved in the access of information about individual objects (Tyler et al., 2004, 2013; Moss et al., 2005; Clarke et al., 2011), as well as specific people and places (Tranel, 2006; Drane et al., 2009).

In this study, we used analyses of multivoxel pattern information to identify both similarities and differences in the semantic processing of words and objects. Data-driven RSA methods revealed key differences in the representational topography across stimulus modalities, even in regions that were sensitive to semantic category representations in both modalities. Specifically, Lp-MTG reflects semantic information across modalities, but this information takes a modality-specific form, whereas the results for LIPS suggest that it is involved in modality-invariant targeted retrieval of task-relevant semantic feature information. These results go beyond identifying individual regions involved in amodal processing and show how the computational properties of a region vary as a function of the network in which it is engaged.

References

Akama H, Murphy B, Na L, Shimizu Y, Poesio M (2012) Decoding semantics across fMRI sessions with different stimulus modalities: a practical MVPA study. Front Neuroinform 6:24. CrossRef Medline

Baayen R, Piepenbrock R, van Rijn H (1993) The CELEX lexical database on CD-ROM. Philadelphia: Linguistic Data Consortium.

Badre D, Poldrack RA, Paré-Blagoev EJ, Insler RZ, Wagner AD (2005) Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. Neuron 47:907–918. CrossRef Medline

Bates E, D'Amico S, Jacobsen T, Székely A, Andonova E, Devescovi A, Herron D, Lu CC, Pechmann T, Pléh C, Wicha N, Federmeier K, Gerdjikova I, Gutierrez G, Hung D, Hsu J, Iyer G, Kohnert K, Mehotcheva T, Orozco-

Figueroa A, Tzeng A, Tzeng O (2003) Timed picture naming in seven languages. Psychon Bull Rev 10:344–380. CrossRef Medline

Beauchamp MS, Lee KE, Haxby JV, Martin A (2003) fMRI responses to video and point-light displays of moving humans and manipulable objects. J Cogn Neurosci 15:991–1001. CrossRef Medline

Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 19:2767–2796. CrossRef Medline

Binney RJ, Embleton KV, Jefferies E, Parker GJM, Ralph MAL (2010) The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: evidence from a novel direct comparison of distortion-corrected fMRI, rTMS, and semantic dementia. Cereb Cortex 20:2728–2738. CrossRef Medline

Bouchard KE, Mesgarani N, Johnson K, Chang EF (2013) Functional organization of human sensorimotor cortex for speech articulation. Nature 495:327–332. CrossRef Medline

Bright P, Moss H, Tyler LK (2004) Unitary vs multiple semantics: PET studies of word and picture processing. Brain Lang 89:417–432. CrossRef Medline

Bright P, Moss HE, Longe O, Stamatakis EA, Tyler LK (2007) Conceptual structure modulates anteromedial temporal involvement in processing verbally presented object properties. Cereb Cortex 17:1066–1073. CrossRef Medline

Cattell RB (1966) The scree test for the number of factors. Multivariate Behav Res 1:245–276. CrossRef

Chao LL, Haxby JV, Martin A (1999) Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat Neurosci 2:913–919. CrossRef Medline

Chee MWL, Weekes B, Lee KM, Soon CS, Schreiber A, Hoon JJ, Chee M (2000) Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. Neuroimage 12:392–403. CrossRef Medline

Clarke A, Taylor KI, Tyler LK (2011) The evolution of meaning: spatiotemporal dynamics of visual object recognition. J Cogn Neurosci 23:1887–1899. CrossRef Medline

Connolly AC, Guntupalli JS, Gors J, Hanke M, Halchenko YO, Wu YC, Abdi H, Haxby JV (2012) The representation of biological classes in the human brain. J Neurosci 32:2608–2618. CrossRef Medline

Creem-Regehr SH, Lee JN (2005) Neural representations of graspable objects: are tools special? Brain Res Cogn Brain Res 22:457–469. CrossRef Medline

Creem-Regehr SH, Dilda V, Vicchrilli AE, Federer F, Lee JN (2007) The influence of complex action knowledge on representations of novel graspable objects: evidence from functional magnetic resonance imaging. J Int Neuropsychol Soc 13:1009–1020. CrossRef Medline

Cristescu TC, Devlin JT, Nobre AC (2006) Orienting attention to semantic categories. Neuroimage 33:1178–1187. CrossRef Medline

Drane DL, Ojemann GA, Ojemann JG, Aylward E, Silbergeld DL, Miller JW, Tranel D (2009) Category-specific recognition and naming deficits following resection of a right anterior temporal lobe tumor in a patient with atypical language lateralization. Cortex 45:630–640. CrossRef Medline

Dronkers NF, Wilkins DP, Van Valin RD Jr, Redfern BB, Jaeger JJ (2004) Lesion analysis of the brain areas involved in language comprehension. Cognition 92:145–177. CrossRef Medline

Fairhall SL, Caramazza A (2013) Brain regions that represent amodal conceptual knowledge. J Neurosci 33:10552–10558. CrossRef Medline

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1:1–47. CrossRef Medline

Guenther FH, Ghosh SS, Tourville JA (2006) Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang 96:280–301. CrossRef Medline

Hoeren M, Kaller CP, Glauche V, Vry MS, Rijntjes M, Hamzei F, Weiller C (2013) Action semantics and movement characteristics engage distinct processing streams during the observation of tool use. Exp Brain Res 229:243–260. CrossRef Medline

Holland R, Lambon Ralph MA (2010) The anterior temporal lobe semantic hub is a part of the language neural network: selective disruption of irregular past tense verbs by rTMS. Cereb Cortex 20:2771–2775. CrossRef Medline

Humphries C, Binder JR, Medler DA, Liebenthal E (2007) Time course of semantic processes during sentence comprehension: an fMRI study. Neuroimage 36:924–932. CrossRef Medline

Kriegeskorte N, Kievit RA (2013) Representational geometry: integrating cognition, computation, and the brain. Trends Cogn Sci (Regul Ed) 17: 401–412. CrossRef Medline

Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. Proc Natl Acad Sci U S A 103:3863–3868. CrossRef Medline

Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis—connecting the branches of systems neuroscience. Front Syst Neurosci 2:4. CrossRef Medline

Lau EF, Phillips C, Poeppel D (2008) A cortical network for semantics: (de)constructing the N400. Nat Rev Neurosci 9:920–933. CrossRef Medline

Mahon BZ, Milleville SC, Negri GAL, Rumiati RI, Caramazza A, Martin A (2007) Action-related properties shape object representations in the ventral stream. Neuron 55:507–520. CrossRef Medline

Mahon BZ, Schwarzbach J, Caramazza A (2010) The representation of tools in left parietal cortex is independent of visual experience. Psychol Sci 21:764–771. CrossRef Medline

Martin A (2007) The representation of object concepts in the brain. Annu Rev Psychol 58:25–45. CrossRef Medline

Meyer K, Kaplan JT, Essex R, Webber C, Damasio H, Damasio A (2010) Predicting visual stimuli on the basis of activity in auditory cortices. Nat Neurosci 13:667–668. CrossRef Medline

Moss HE, Rodd JM, Stamatakis EA, Bright P, Tyler LK (2005) Anteromedial temporal cortex supports fine-grained differentiation among objects. Cereb Cortex 15:616–627. CrossRef Medline

Mur M, Bandettini PA, Kriegeskorte N (2009) Revealing representational content with pattern-information fMRI—an introductory guide. Soc Cogn Affect Neurosci 4:101–109. CrossRef Medline

Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum Brain Mapp 15:1–25. CrossRef Medline

Noonan KA, Jefferies E, Visser M, Lambon Ralph MA (2013) Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. J Cogn Neurosci 25:1824–1850. CrossRef Medline

Patterson K, Nestor PJ, Rogers TT (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci 8:976–987. CrossRef Medline

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830.

Rissman J, Gazzaley A, D'Esposito M (2004) Measuring functional connectivity during distinct stages of a cognitive task. Neuroimage 23:752–763. CrossRef Medline

Rizzolatti G, Matelli M (2003) Two different streams form the dorsal visual system: anatomy and functions. Exp Brain Res 153:146–157. CrossRef Medline

Shinkareva SV, Malave VL, Mason RA, Mitchell TM, Just MA (2011) Commonality of neural representations of words and pictures. Neuroimage 54:2418–2425. CrossRef Medline

Simanova I, Hagoort P, Oostenveld R, Gerven MAJ van (2012) Modality-independent decoding of semantic information from the human brain. Cereb Cortex. Advance online publication. Retrieved December 7, 2013. doi:10.1093/cercor/bhs324. CrossRef Medline

Snodgrass JG, Corwin J (1988) Perceptual identification thresholds for 150 fragmented pictures from the Snodgrass and Vanderwart picture set. Percept Motor Skills 67:3–36. CrossRef Medline

Székely A, Bates E (2000) Objective visual complexity as a variable in studies of picture naming, Vol 12, number 2. La Jolla, CA: Center for Research in Language, University of California, San Diego.

Taylor KI, Moss HE, Stamatakis EA, Tyler LK (2006) Binding crossmodal object features in perirhinal cortex. Proc Natl Acad Sci U S A 103:8239–8244. CrossRef Medline

Taylor KI, Devereux BJ, Acres K, Randall B, Tyler LK (2012) Contrasting effects of feature-based statistics on the categorisation and basic-level identification of visual objects. Cognition 122:363–374. CrossRef Medline

Tranel D (2006) Impaired naming of unique landmarks is associated with left temporal polar damage. Neuropsychology 20:1–10. CrossRef Medline

Tyler LK, Marslen-Wilson W (2008) Fronto-temporal brain systems supporting spoken language comprehension. Philos Trans R Soc Lond B Biol Sci 363:1037–1054. CrossRef Medline

Tyler LK, Stamatakis EA, Bright P, Acres K, Abdallah S, Rodd JM, Moss HE (2004) Processing objects at different levels of specificity. J Cogn Neurosci 16:351–362. CrossRef Medline

Tyler LK, Cheung TPL, Devereux BJ, Clarke A (2013) Syntactic computations in the language network: characterizing dynamic network properties using representational similarity analysis. Front Psychol 4:271. CrossRef Medline

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15:273–289. CrossRef Medline

Valyear KF, Cavina-Pratesi C, Stiglick AJ, Culham JC (2007) Does tool-related fMRI activity within the intraparietal sulcus reflect the plan to grasp? Neuroimage 36 [Suppl 2]:T94–T108. CrossRef

Vandenberghe R, Price C, Wise R, Josephs O, Frackowiak RSJ (1996) Functional anatomy of a common semantic system for words and pictures. Nature 383:254–256. CrossRef Medline

Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, Mazoyer B, Tzourio-Mazoyer N (2006) Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. Neuroimage 30:1414–1432. CrossRef Medline

Visser M, Jefferies E, Embleton KV, Lambon Ralph MA (2012) Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. J Cogn Neurosci 24:1766–1778. CrossRef Medline

Warrington EK, Shallice T (1984) Category specific semantic impairments. Brain 107:829–854. CrossRef Medline