

## Does the sound of a barking dog activate its corresponding visual form?

Reilly, Jamie; Garcia, Amanda; Binney, Richard J

### Brain and Language

DOI:

[10.1016/j.bandl.2016.05.006](https://doi.org/10.1016/j.bandl.2016.05.006)

Published: 01/08/2016

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

*Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):*

Reilly, J., Garcia, A., & Binney, R. J. (2016). Does the sound of a barking dog activate its corresponding visual form? An fMRI investigation of modality-specific semantic access. *Brain and Language*, 159, 45-59. <https://doi.org/10.1016/j.bandl.2016.05.006>

#### Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Running head: MODALITY-SPECIFIC SEMANTIC ACCESS

**Does the Sound of a Barking Dog Activate its Corresponding Visual Form?  
An fMRI Investigation of Modality-Specific Semantic Access**

Jamie Reilly<sup>1,2,\*</sup>, Amanda Garcia<sup>3</sup>, and Richard J. Binney<sup>1,2</sup>

<sup>1</sup>Eleanor M. Saffran Center for Cognitive Neuroscience

<sup>2</sup>Department of Communication Sciences and Disorders  
Temple University, Philadelphia, PA USA

<sup>3</sup>Department of Clinical and Health Psychology  
University of Florida, Gainesville, FL USA

\*Corresponding author:

Jamie Reilly, Ph.D.

Temple University

Weiss Hall

Philadelphia, PA 19122, USA

Email: [reillyj@temple.edu](mailto:reillyj@temple.edu)

Tel: (215) 204-3995

**Abstract**

Much remains to be learned about the neural architecture underlying word meaning. Fully distributed models of semantic memory predict that the sound of a barking dog will conjointly engage a network of distributed sensorimotor spokes. An alternative framework holds that modality-specific features additionally converge within transmodal hubs. Participants underwent functional MRI while covertly naming familiar objects versus newly learned novel objects from only one of their constituent semantic features (visual form, characteristic sound, or point-light motion representation). Relative to the novel object baseline, familiar concepts elicited greater activation within association regions specific to that presentation modality. Furthermore, visual form elicited activation within high-level auditory association cortex. Conversely, environmental sounds elicited activation in regions proximal to visual association cortex. Both conditions commonly engaged a putative hub region within lateral anterior temporal cortex. These results support hybrid semantic models in which local hubs and distributed spokes are dually engaged in service of semantic memory.

**Keywords:** Semantic cognition; Anterior Temporal Lobe; Semantic Access; Concept Acquisition; Naming

## 1. Introduction

The *binding problem* reflects a longstanding question within the cognitive neurosciences regarding the manner(s) in which the human brain amalgamates features from disparate sensorimotor modalities into coherent conceptual representations. Object knowledge is comprised of complex conjunctions of features from visual, auditory, olfactory, haptic and motoric modalities, in addition to a range of other affective, episodic and lexical associations. Much of the early stage perceptual processing associated with these modalities occurs within distinct and remote regions of the brain. Accordingly, there exists a rich and long-standing debate regarding the mechanistic neural architecture that binds multimodal feature information into coherent wholes.

Perhaps the most historically dominant neurocognitive hypothesis of conceptualization is that it reflects auto-associated and conjoint activation of information stored in the multiple remote modality-specific regions of the brain (Wernicke, 1874; Hauk, Johnsrude, & Pulvermüller, 2004; Eggert, 1977; Gage & Hickok, 2005; Pulvermüller, 2001). For example, hearing the distinct ‘meow’ of a cat will activate auditory association cortices in addition to a range of associated features in other modalities (e.g., form, color) (See Figure 1, Panel 1). There is considerable evidence for these distributed-only accounts in the form of functional neuroimaging studies that have demonstrated that different types of concept (e.g., action concepts vs. object concepts) activate modality-specific association cortices (motor cortex vs. visual cortex) linked to their most crucial sources of defining information (motor sequences to perform them vs. the way they look) (Hoenig, Sim, Bochev, Herrnberger, & Kiefer, 2008; Kellenbach, Brett, & Patterson, 2001; Simmons, Martin, & Barsalou, 2005; Trumpp, Kliese, Hoenig, Haarmeier, & Kiefer, 2012).

--Insert Figure 1 about here--

Alternative perspectives argue the necessity for transmodal brain regions; tertiary cortical areas that transcend particular modalities. It has been hypothesized that one or more of these regions fulfill a central organizing principle by which multiple modalities can be drawn together for the process of conceptualization (see Figure 1, Panel 2). In particular, such a region constitutes a convergence zone that is massively reciprocally linked to the distributed network of sensorimotor regions and thus uniquely suited as a nexus point for coordination of polymodal feature processing (Damasio, 1989). It has been further proposed that such ‘hubs’ play a necessary role in the computations required for coherent conceptual representation (Lambon Ralph, Sage, Jones, & Mayberry, 2010; Rogers & McClelland, 2004).

Patterns of behavioral performance among patients with relatively focal versus diffuse neuropathologies present a compelling test of these competing hypotheses. A theoretical model with no center (i.e., distributed only) predicts that only catastrophic, diffuse, bilateral brain damage is sufficient to cause multi-modal semantic impairment. In contrast, the hub perspective holds that profound semantic memory disorders can result from relatively circumscribed damage focused upon regions of transmodal representational cortex. The case for the existence of a transmodal hub for conceptual representation has historically relied heavily on studies of patients with semantic dementia. This clinical population exhibits progressive, yet relatively circumscribed bilateral anterior temporal atrophy coupled with a selective dissolution of conceptual knowledge that transcends task and modality (Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000; Coccia, Bartolini, Luzzi, Provinciali, & Lambon Ralph, 2004; Goll et al., 2010; Piwnica-Worms, Omar, Hailstone, & Warren, 2010). Decades worth of detailed neuropsychological investigations of semantic dementia provide a compelling foundation of support for the hypothesis that the bilateral anterior temporal lobes play key roles in computing transmodal

conceptual representations (Lambon Ralph, 2014; Patterson, Nestor, & Rogers, 2007). These assertions primarily informed by patient-based dissociations have been bolstered by a growing body of convergent evidence from functional imaging and neurostimulation studies (Binney, Embleton, Jefferies, Parker, & Lambon Ralph, 2010; Halgren et al., 2006; Marinkovic et al., 2003; Mion et al., 2010; Pobric, Jefferies, & Lambon Ralph, 2007; Shimotake et al., 2014; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996; Visser & Lambon Ralph, 2011).

Evidence for both distributed and local (hub) components of the neural architecture subserving semantic cognition is reconciled within a class of hybrid, pluralistic theories, the most influential, to date, being the Hub and Spoke model proposed by Patterson, Lambon Ralph, Rogers and colleagues (Hoffman, Jones, & Ralph, 2012; Lambon Ralph & Patterson, 2008; Rogers & McClelland, 2004). The architecture and functioning of the Hub and Spoke model is made explicit by the connectionist computational implementation of Rogers et al. (Rogers et al., 2004; see also Patterson et al., 2007), wherein modality-specific spokes are interfaced via a central transmodal hub. According to Lambon Ralph (2014) and colleagues, the spokes are the substrate of invariant representations important for recognition of perceptual objects within their respective modality and the coding of similarities between different perceptual objects. The surface similarities computed by the spokes are necessary but insufficient for the formation of coherent and generalizable concepts, providing only a fragmentary guide to meaning. Herein lies the importance of a hub. The additional representational layer afforded by the ATL hub provides a tertiary level of abstraction allowing for distillation of the highly complex, non-linear transmodal relationships between multi-modal features that comprise concepts (Lambon Ralph et al., 2010; Reilly, Peelle, Garcia, & Crutch, 2016). Importantly, this does not imply that conceptualization can be achieved solely by activation of the representations subserved by the hub. Instead, in the Rogers *et al.* model, the hub and spokes are bi-directionally connected and complete

conceptualization arises from the conjoint action of both the transmodal hub and each of the modality specific sources of information (Pobric, Jefferies & Lambon Ralph, 2010). This approach, therefore, predicts that conceptual processing will be reflected in activation of both transmodal representational cortex (in particular, the ATL) and the distributed network of modality-specific association regions, and that this full network will be activated regardless of the input modality through which the concept is probed. Neuroimaging techniques such as fMRI are useful tools to explore this hypothesis, given they enable visualization of activation across the whole brain when individuals are performing semantic tasks. Several prior fMRI and PET studies have compared activation associated with conceptual processing probed in different modalities (Taylor, Moss, Stamatakis, & Tyler, 2006; Thierry & Price, 2006; Tranel, Grabowski, Lyon, & Damasio, 2005) or probed via verbal versus non-verbal (pictures) representational formats (Jouen et al., 2014; Thierry & Price, 2006; Vandenberghe et al., 1996; Visser & Lambon Ralph, 2011). However, the prevailing emphasis in such studies has been upon identifying regions that are commonly activated across modalities in hope of identifying putative transmodal ‘hub’ regions (Binder, Desai, Graves, & Conant, 2009). Less attention has been paid to the contribution of spoke regions. Indeed, there exists no clear operationalization of what constitutes a spoke region within the brain. As far as we are aware, there have been no explicit predictions under the hub and spoke framework beyond that they are higher-order modality-specific association regions that lie at the apex of each unimodal processing streams. Some predictions could be derived from a parallel set of functional imaging studies motivated by the distributed-only perspective that, as discussed above, holds that conceptual processing recruits the same modality-specific regions involved in perception and action. Such studies have identified lateral and ventral temporo-occipital regions that show responses selective to processing of visual features of objects and object nouns, such as visual form, color and motion, and that certain sub-areas of these brain regions preferentially respond to certain semantic categories of

objects such as faces, animals and tools (Chao, Haxby, & Martin, 1999; Chao, Weisberg, & Martin, 2002; Kanwisher & Yovel, 2006; Wheatley, Weisberg, Beauchamp, & Martin, 2005). Other studies have found similarly selective profiles of regions of the lateral temporal cortex for auditory features (e.g., environmental sounds) (Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008), the ventral premotor cortex for object manipulability (Chao & Martin, 2000), and the orbito-frontal cortex for olfactory and gustatory features (Goldberg, Perfetti, & Schneider, 2006). There is also complementary evidence for a contribution of distributed sensorimotor association regions to concepts stemming from neuropsychological investigations and electrophysiological and neurostimulation studies (Reilly et al., 2011; Kemmerer, Castillo, Talavage, Patterson, & Wiley, 2008; Pulvermüller, 2013; Pulvermüller, Shtyrov, & Ilmoniemi, 2005; Shtyrov & Pulvermüller, 2007; Pobric et al., 2010). In the present study, we sought to investigate whether some of these potential candidate ‘spoke’ regions (See Methods) would show conjoint activation with the ATL hub region during a task involving semantic processing, and whether the differential roles of the hub and spoke regions can be observed by virtue of their response profiles. In particular, spoke regions should show an effect of modality in line with their unimodal function, but we were specifically interested in whether they exhibit a differential response according to a semantic manipulation which would indicate that they are involved in conceptual processing. Moreover, in line with the connectionist implementation of the hub and spoke model, we hypothesized that the spoke regions would show a semantic effect on their activation regardless of the modality of stimulus presentation. Regarding the ATL hub, we hypothesized an effect of the semantic manipulation but no effect of modality (i.e., all representational modalities ultimately converge upon an amodal semantic representation) in line with its purported transmodal function.

To the best of our knowledge, there have been no prior functional imaging studies that have attempted to simultaneously test hypotheses regarding interactivity between hub and spoke regions in

this manner. This could reflect, at least in part, the considerable methodological challenge of isolating activation specific to semantic processes from that associated with sensorimotor perceptual processes per se. Conventional subtraction analyses in fMRI typically contrast activation associated with semantic tasks (unimodal or crossmodal) to that associated with non-semantic perceptual judgments (unimodal or cross-modal) in order to subtract out low-level perceptual processes or cross-modal integration per se (i.e., at a pre-semantic stage) and control for domain-general function such as decision-making processes. The baseline control tasks often use stimuli such as false fonts or noise vocoded speech that have been distorted or scrambled to remove their ‘meaningfulness’ but retain, at least partially, perceptual complexity and the decision-making element of the task. However, the complexity of stimuli and the task difficulty often remain variable between the semantic (experimental) and control (low level perceptual baseline) conditions and thus semantic processes continue to be confounded with other processes in the interpretation of activations. Moreover, control conditions for different modalities may vary in their effectiveness to control for low-level processes and, as such, comparisons between activations in one modality and another are difficult to interpret. Alternative approaches include parametric modulation of the ‘meaningfulness’ of stimuli, to which activation related to semantic processing should covary. We undertook a novel approach to addressing the role of ‘spoke’ regions in semantic representation by contrasting activation when subjects named familiar entities compared to newly learned object concepts. Our rationale for choosing newly learned objects as a baseline was that the familiar and novel stimuli would be well matched in terms of perceptual complexity and that task requirements would be identical for the two stimulus sets. Moreover, the familiar/novel exemplar distinction reflects a manipulation of ‘meaningfulness’ that may be used to yield activations associated with evocation of conceptual representations. The novel entities were fictional animals and artefacts with distinct form, motion and sounds. Thus, these objects might be considered meaningful novel

exemplars of superordinate categories (e.g., animals) but with relatively impoverished representation as basic level concepts or unique exemplars/entities as they do not have the multiple episodic, affective or encyclopedic associations (to name a few) that enrich the conceptualization of familiar animals/artefacts. Contemplating connectionist frameworks, we hypothesized that the familiar concepts will have stronger connection weightings and more robust neural representations, and thus would evoke more robust activation across the neural network subserving conceptual representation. Under this assumption we attempted to use this contrast (familiar > novel) to yield semantic activations while controlling for pre- and post-semantic processes/activations. We examined activations associated with semantic processing when the input was constrained to one of three feature types (an object's static visual form, point light motion or characteristic sound) which each have dissociable cortical regions associated with perceptual processing (See Methods). We used whole brain and targeted region of interest analyses to reveal differences and commonalities between the networks activated by conceptual processing probed by different stimulus presentation modalities.

## **2. Method**

### **2.1. Study Design**

We examined patterns of activation for a series of familiar object concepts (i.e., bird, dog, scissors, clock) relative to four newly learned object concepts (plufky, korbok, wilzig, blerga). During the learning phase, participants learned the names of these novel concepts by watching animated videos of the target concepts in action, moving in distinctive paths/manners and making unique animal/tool noises while a narrated voiceover announced each item's name. Three days later the participants underwent the same exposure condition. One week after initiation of the learning phase, we scanned

participants as they repeatedly named novel and familiar items upon exposure to only one of their modality-specific semantic features. We segregated the three input modalities (i.e., visual form, environmental sound, and point light motion) into separate scanning runs. For example, in the visual form session participants named DOG from a grayscale photograph. In the auditory session, participants named DOG from the sound of a barking dog, and in the motion session participants named DOG from a point-light video depicting the canonical motion of a dog walking. Our analyses were aimed at detecting activations both associated to each input modality and areas common across all the modalities. In contrasting familiar versus novel items, we aimed to isolate deep semantic processing via a cognitive subtraction. Novel object are matched to familiar in terms of perceptual processing (items were approximately comparable in visual complexity; see Figure 2), and phonological encoding demands. Training of novel items was continued until 100% accuracy was achieved such that differences in activations for novel and familiar items was unlikely to be attributed to difficulty. As such, the familiar-novel contrast was intended to isolate enhanced activation associated with retrieval of conceptual knowledge associated with familiar objects.

## **2.2. Participants**

18 healthy young adults (14 females) were recruited from the University of Florida. Participants reported no prior neurological injury, learning disability (e.g., dyslexia), and were not on a current regimen of sedative medications. All were right-hand dominant, native English speakers. Mean age was 22.8 and mean education was 16.1 years. Participants provided informed consent in accord with the Institutional Review Board of the University of Florida and were nominally compensated for taking part in the study.

## **2.3. Stimuli**

The familiar objects included two animals (dog, bird) and two manufactured artifacts (clock, scissors). Items from both the living and non-living domain were included such that our results were not entangled with category-specific effects. Visual form presentation involved grayscale, cartoon-like photographs of the target items on a black background. Auditory stimuli included clips (e.g., ticking clock) originally obtained from online sound repositories which we subsequently trimmed to 2000ms, matched in sampling rate and normalized on amplitude using the GoldWave waveform editor. The motion stimuli involved 2000ms videos of white points placed along articulated joints and edges against a black background.

Four novel objects were learned one week before scanning. Participants learned these names by watching 20s long videos of the target objects moving in a distinctive path/manner and making a distinct sound while a male narrator announced each item's name. Novel sounds were created by first subjecting real sounds to low-pass filtering and warbling effects using the GoldWave waveform editor. The audio clips were then all matched in duration and batch normalized for root mean square amplitude (i.e., perceived as roughly equal volume). Point light videos analogous to the familiar items were created using LightWave animation software. Figure 2 illustrates the stimuli characteristics.

--Insert Figure 2 about here--

## **2.4. Procedure**

### **2.4.1. Learning phase**

On Day 1 of the experiment, participants learned the names of four novel concepts by watching videos depicting simultaneous motion, sound, and form. Participants passively viewed each video two times, and the order of video presentation was randomized. Upon completion of this learning sequence,

we tested naming ability by showing truncated (2s) clips of each target item and asking the participant to name the item. For any participant who failed to name a single item, we repeated presentation of the entire video sequence (4 items \* 2 repetitions) and probed naming until that participant achieved 100% accuracy. The learning session was terminated when participants named all items with 100% accuracy. We repeated this session using the exact parameters three days later.

#### **2.4.2. Imaging phase**

On the day of the scanning session, participants were informed that they were to name the items they had learned a week before in addition to several untrained but familiar items but that only one of their constituent features would be presented in a given scanning run. Furthermore, they were instructed that overt naming was required on only a subset of trials and this would be cued at random during the experiment. Before entering the scanner, we conducted a brief (5 minute) familiarization procedure where participants performed the task with experimenter feedback in each of the test modalities (point light, sound, picture). Again, 100% accuracy for naming familiar and novel items was required before participants proceeded to scanning.

Participants were instructed to overtly name target stimuli as quickly and accurately as possible from the onset of the cue (a tilde presented in isolation following stimulus presentation). Probes were pseudorandomly interspersed over the duration of a scanning run, with 24 of 64 trials requiring an overt speech response. As such, participants were unaware on any given trial of whether they were required to make an overt response, and as such were encouraged to identify the object on each stimulus presentation (covert naming).

Participants completed three functional scanning runs (8 minutes each) spaced over one hour, with approximately 2-4 minutes of inactivity between each. Each functional scanning run contained only

features presented in one of the modalities (i.e., visual form, sound, or point-light motion). Each object was presented 8 times for a duration  $\approx 2000$ ms, for a total of 64 pseudo-randomized stimulus presentations per scanning run. Each object was cued for naming 3 times per scanning run. The intervals between stimulus presentation and onset of the cue was 2s. An interstimulus interval of either 2s or 4s was pseudorandomly assigned across stimuli. Figure 3 illustrates the trial structure and timing parameters.

--Insert Figure 3 about here--

We presented stimuli using a laptop computer running E-Prime 2.0 Professional software. We synchronized scanning with stimulus presentation by timelocking the offset of each trial to the radiofrequency (RF) pulse to the using a TTL synchronization box (Nordic NeuroScan, Inc). Trial durations were a maximum of 2000ms (TR duration) but sometimes slightly shorter due to receipt of the pulse signal. This procedure ensured that no cumulative timing errors were introduced into the experiment as a result of stimulus buffering.

Participants viewed all picture and motion stimuli on a monitor situated behind the scanner bore via a mirror slotted onto the head coil. Participants heard auditory stimuli over MR compatible pneumatic headphones (ScanSound Inc).

## **2.5. Imaging acquisition**

All imaging was performed on a Philips Achieva 3 Tesla scanner with a 32-channel SENSE head coil. We acquired a high resolution T1-weighted anatomical image using a magnetization-prepared rapid acquisition with gradient echo (MPRAGE) sequence with the following parameters: in-plane resolution

= mm<sup>2</sup>, Field of View (FOV)=240mm<sup>2</sup>, slice thickness=1mm, Number of Slices=170, Flip angle=8°; Repetition Time (TR)=7.0ms; Echo Time (TE)=3.2ms. The gradient-echo echo-planar fMRI sequences were performed with the following parameters; in-plane resolution =3mm x 3mm, Field of View (FOV)=240mm<sup>2</sup>, Slice thickness=4mm, Number of Slices=38, Flip angle=90°; Repetition Time (TR)=2000ms; Echo Time (TE)=30ms, Slice Gap=None; Interleaved Slice Acquisition, Number of Dynamic Scans=242. We restricted FOV for temporal and inferior parietal lobe coverage, excluding the highest convexities of the dorsal cerebrum.

## 2.6. fMRI Data Analysis

Analysis was carried out using statistical parametric mapping (SPM8) software (Wellcome Trust Centre of Neuroimaging, UK). Within each subject, the functional EPI volumes from across all three scanning sessions were re-aligned using a 6-parameter rigid body transform estimated using a least squares approach and a two-pass procedure. The aligned images and a mean volume were then re-sliced using fourth-degree B-spline interpolation. This procedure corrected for differences in subject positioning between sessions and minor motion artefacts within a session. The T1-weighted anatomical image was registered to the mean functional volume using a six-parameter rigid-body transform and subsequently subjected to SPM8's unified tissue segmentation and normalization procedure in order to estimate a spatial transformation from subject space to a standard stereotactic space, according to the Montreal Neurological Institute (MNI) protocol, for inter-subject averaging. This transform was applied to each of the subject's functional volumes, resampling to a 3 x 3 x 3 mm voxel size using trilinear interpolation and preserving the intensities of the original images (i.e., no "modulation"). These volumes were then smoothed with an 8 mm full-width half-maximum (FWHM) Gaussian filter.

Data were analyzed using a general linear model approach (GLM). At the individual subject level, each scanning run (and thus presentation modality) was modeled within a separate fixed effects analysis. Our rationale for analyzing each run/modality in a separate model was that some participants did not complete all three runs/modalities due to technical difficulties with audio presentation. Within the model for a given scanning run, the presentation of familiar items (e.g., Dog) and newly learned items (e.g., Plufky) were modelled as two separate boxcar functions. Instances where subjects made an overt verbal response (probed trials only) were also modelled with a two-second boxcar function in order to capture speech-related changes in BOLD as an independent nuisance covariate. Rest periods were modelled implicitly. A set of six motion parameters for each scan session, which were estimated during the realignment step, were also included as nuisance regressors. Regressors were subsequently convolved with a canonical hemodynamic response function. Data were further treated with a high-pass filter with a cutoff of 128 s. Planned contrasts were calculated to assess differences in activation between naming familiar items and newly-learned items [Familiar – Novel]. These contrast images were entered into subsequent multi-subject mixed-effect analyses directed at our a priori hypotheses, as follows.

## **2.7. Whole-brain analyses of modality-specific effects**

First, we examined activation for semantic access of familiar entities relative to novel entities in separate whole-brain analyses for each of the stimulus presentation modalities (one sample t-tests). This allowed for an initial assessment of brain-wide similarities and differences in the topology of activations recruited by semantic processing of stimuli in the different modalities. Statistical parametric maps were thresholded with an uncorrected voxel-height threshold of  $P < 0.005$  and a minimum cluster size of 30 contiguous voxels. Inferences were made on activations surviving this threshold if they occurred in a

priori predicted regions. These predicted regions included occipito-temporal visual association cortex and superior temporal auditory association cortex and also regions implicated in the semantic neuroscience literature, namely, the anterior temporal lobe, the posterolateral temporal lobe, the frontal operculum and ventral parietal cortex. Activations outside of these predicted regions were assessed using a more stringent threshold corrected for the multiple comparisons problem using the topological false discovery rate as implemented in SPM8 ( $p < 0.05$ ). All reported coordinates are in Montreal Neurological Institute (MNI) space.

## **2.8. Whole-brain analyses of common activation to familiar > novel entities across different modalities**

Our second planned analysis sought to identify regions areas activated by semantic processing irrespective of the feature type (form, sound or motion) that was presented, and thus fulfill hypothesized characteristics of high-order regions involved in transmodal semantic processing, such as the ATL hub. In particular, we planned a conjunction analysis (within a one-way Analysis of Variance (ANOVA)) targeting voxels that are significantly activated independently in both conditions/modalities (Nichols et al, 2005). The threshold for this analysis was  $p < 0.01$ , uncorrected, in effect being 0.001 for pairwise commonalities. We also performed post hoc analyses assessing main effects and interactions of modality and familiarity within a factorial ANOVA. This required calculation of contrasts at the first level for [Visual Familiar] only, [Visual Novel] only, [Auditory Familiar] only, and so on. Statistical parametric maps generated by these analysis were assessed following an application of a  $p < 0.001$ , uncorrected voxel-height threshold and an FDR-corrected cluster extent threshold at  $p < 0.05$ .

## **2.9. Targeted regions of interest (ROI) analyses**

In addition to the more conservative whole-brain analysis, we applied a planned targeted region of interest (ROI) analysis using the MARSBAR toolbox (Brett, Anton, Valabregue, & Poline, 2002). This method is an alternative to approach to overcoming the multiple comparisons problem and is ideal for regional hypothesis testing. We reasoned that differences in activation to familiar relative to novel concepts may be subtle and thus less conservative analyses may be required to detect them. Moreover, plotting the parameter estimates extracted from an ROI can provide a more intuitive means to interpreting differential effects compared to statistical maps. MARSBAR calculates a single summary value to represent activation across all voxels within a given ROI (in the present case, the median of the parameter estimates). Details on ROI selection and definition are provided below. We extracted an estimate for the [Familiar-Novel] contrast for each ROI in each modality and performed group analyses using statistical software outside of SPM. In the group level statistics (Figure 6), a positive value indicates greater activation for familiar item, a negative indicates greater activation for novel items, and a zero or insignificant value indicates no differential activation between familiar and novel items.

## **2.10. Region of Interest Definition**

Candidate ‘spoke’ regions of interest (ROIs) were defined on the basis of prior functional imaging data from studies investigating higher order and/or semantic processing of sensorimotor features within unimodal association cortices. Given the presentation modalities we investigated specifically targeted (i) superior temporal (anterior parabelt) auditory association cortex that is associated with processing of higher-order auditory objects, and (ii) the ventral occipito-temporal visual association cortex that is particularly associated with higher-order processing of object form. An ROI for motion-associated regions was omitted from this analysis, as was the motion condition data, due to not finding an effect at the whole-brain-level in the predicted temporal or parietal regions (see above), even

at the more liberal threshold (see Results). A third ROI represented the ATL hub region. Each ROI was spherical with a radius of 10mm (see Figure 6 for a visual depiction).

The precise location of the superior temporal ROI was defined on the basis of work from Warren, Jennings and Griffiths (2005) who identified regions associated with the higher order and abstracted spectral analysis of sounds. We averaged the coordinates of peaks reported by these authors for the contrast they interpreted as relating to an abstraction of spectral shape, which may be relevant to the analysis of auditory sources or objects, such as voices. This average coordinate (57,-22,3) defined the center of mass for the superior temporal ROI. Given the 10mm radius of our ROI, it encompassed many of the individual peak coordinates. We had no prior hypotheses regarding lateralization of activation in this region, and thus included a left hemisphere homologue (-57,-22,3) ROI in the analysis.

It is well established that the ventral occipito-temporal cortex (vOTC) is involved in the processing of higher-level visual properties of objects such as global shape/form. Moreover, it has been demonstrated that portions of the mid-to-posterior fusiform gyrus preferentially respond to certain categories of objects, such as animals or tools (Chao et al., 1999; Chao & Martin, 2000; Wheatley et al., 2005), implying at least a near-conceptual level of processing but still within modality-specific cortex. These results are usually discussed as fitting the predictions of the distributed-only perspective, but they also suggest this region is a good candidate for a visual form ‘spoke’ in the context of the hub-and-spoke model. We used reported coordinates from the fMRI study of Chao et al. (1999) to define our ventral occipito-temporal cortex ROI. Given we were not interested in distinction between categories in the present study, we averaged their coordinates for the animal and tool selective regions (transformed from Talairach coordinates to MNI coordinates using the tal2icbm function; [http://biad02.uthscsa.edu/icbm2tal/tal2icbm\\_spm.m](http://biad02.uthscsa.edu/icbm2tal/tal2icbm_spm.m)). We included both left (-34 -56 -18) and right (35 -53 -20) hemisphere vOTC ROIs.

Recent refinement of hypotheses regarding ATL involvement in conceptual knowledge representation have pinpointed the ventral surface in particular as a transmodal substrate (Binney et al., 2010; Mion et al., 2010; Tyler et al., 2013). Conventional gradient-echo EPI imaging protocols, such as that used in the present study, suffer from signal loss and distortion in this region (Devlin et al., 2000; Embleton, Haroon, Morris, Ralph, & Parker, 2010) and, therefore, reliable BOLD measurements can be difficult to achieve. However, it is possible to obtain good signal within lateral ATL regions that have been equally implicated, bilaterally, on the basis of fMRI and TMS studies, albeit more variably and as function of task demands or input/output modality (Binder, Desai, Graves, & Conant, 2009; Binney et al., 2010; Hoffman, Binney & Lambon Ralph, 2015; Lambon Ralph, Pobric, & Jefferies, 2009; Zahn et al., 2007). We calculated the average of the two sets of coordinates reported for the left lateral ATL in Binney et al.'s (2010) study and used this (-54,6,-28) and its right hemisphere homologue (54,6,-28) to define the center of mass for two ATL ROIs for the present study.

### 3. Results

We acquired full datasets for 13 of 18 participants. For the remaining five participants, the auditory run failed due to an amplifier malfunction. Thus, in the analyses to follow we modeled activation via an unbalanced design including the visual form and point-light motion runs for all 18 participants and auditory runs for 13 participants.

--Insert Figure 4 and Table 1 about here--

#### 3.1. Whole-brain analyses of modality-specific effects

##### 3.1.1. Activation associated with visual form for Familiar relative to Novel Entities [Vis<sub>fam</sub>-Vis<sub>nov</sub>]

Results of the whole brain analysis for familiar >novel entities in the visual form condition are displayed in Figure 4, Row A and Table 1. A large right-hemisphere cluster included three regions associated with transmodal semantic processing and language processing (albeit more so in the left hemisphere), namely the lateral anterior temporal lobe (peaking in the middle temporal gyrus), the inferior frontal gyrus, the posterolateral temporal lobe (including the MTG, STG and STS) and the ventral parietal lobe. This cluster extended ventromedially to reveal greater activation for familiar entities in the posterior and mid parahippocampal gyrus and visual association cortex within the bilateral lingual gyri. As such, familiar concepts invoke greater activation than novel entities in the association cortex tied with the modality of presentation, despite equivalent complexity in the perceptual features of stimuli.

The same cluster extended to auditory association cortex in the anterior superior temporal gyrus and the superior temporal sulcus and also to the posterior planum temporale adjacent to Heschl's gyrus. Therefore, the increased activation for familiar entities extends beyond association cortex for the presentation modality and transmodal association cortex to association cortex of other modalities (in this case, auditory cortex). This large cluster also extended to the right hippocampus, the right insula, the right orbitofrontal cortex and across ventromedial frontal and anterior cingulate cortex bilaterally. A small cluster was also observed in the right thalamus, although it did not survive correction for multiple comparisons.

Four left hemisphere clusters mirrored many of the activations in the right hemisphere cluster. The largest occurred predominately in the left parahippocampal gyrus and left ventral occipitotemporal visual cortex. Another cluster extended over a large swathe of the dorsal posterolateral temporal cortex, ascending to ventral parietal regions and descending to dorsolateral extrastriate visual cortex. A third cluster was observed in left orbitofrontal cortex, while a fourth cluster encompassed portions of left

planum temporale, and left insula, extending toward medial temporal regions. These final two clusters did not survive correction for multiple comparison.

### **3.1.2. Activation associated with environmental Sound for Familiar relative to Novel Entities**

#### **[Sound<sub>fam</sub>-Sound<sub>nov</sub>]**

The familiar > novel entity contrast in the sound condition (Figure 4, Row B and Table 1) revealed a largely symmetrical distribution of activation across the hemispheres. However, large clusters were observed bilaterally at the occipito-temporal-parietal junction encompassing the supramarginal gyri, the posterior superior temporal gyri/sulci and posterior middle temporal gyri. In the right hemisphere, the cluster extended rostrally to auditory cortex in the mid superior temporal gyrus and planum temporale. In the left hemisphere, primary and secondary auditory cortex activation was observed in two smaller clusters, one in the posterior superior temporal gyrus and another more anterior which stretched across the planum temporale to insula cortex. These clusters did not survive correction for multiple comparisons but nevertheless fell within a priori predicted regions.

In both hemispheres, the largest clusters also extended ventrally to extrastriate visual cortex from the posterior middle temporal gyrus. In the right hemisphere, an additional smaller cluster was observed within lateral occipital-temporal visual association cortex with an extension into ventral occipital-temporal regions. This cluster did not survive corrections for multiple comparisons. Contrary to the visual form modality, parahippocampal cortex was not activated.

Three bilateral clusters were observed in posterior cingulate cortex, the precuneus including retrosplenial cortex (this cluster also encroached on the lingual gyri bilaterally) and medial prefrontal cortex.

### 3.1.3. Activation associated with point-light Motion for Familiar relative to Novel Entities

#### [Motion<sub>fam</sub>-Motion<sub>nov</sub>]

In the point-light motion condition, the familiar minus novel contrast revealed two small clusters of activation, (Figure 4, Row C and Table 1). These clusters were located within the ventral central sulcus and ventral postcentral gyrus in the left and right hemisphere, respectively. Neither cluster fell in a priori predicted regions, nor did they survive correction for multiple comparisons. In light of this null result, we excluded the point-light motion condition from subsequent analyses reasoning that it would further produce a null result in a conjunction analysis.

### 3.2. Whole-brain analyses of common activation to familiar > novel entities across the visual form and sound conditions

The next set of analyses continued with a whole-brain approach but sought to identify brain regions activated by semantic processing irrespective of the feature type (visual form or sound) that was presented. In particular, we planned a conjunction analysis to reveal overlap of regions activated in the familiar-novel contrast in each modality. Before we present those findings, however, we will describe the result of a post-hoc analysis in which we used a factorial ANOVA to examine the main effect of Familiarity, as well as that of Modality and the interaction effect.

The results of an F-test on the main effect of Modality are presented in Table 2 and Figure 5, Row A. Modality-specific activations were observed in visual primary and association cortex from the occipital poles to medial occipito-temporal cortex bilaterally and in bilateral auditory cortex including Heschl's gyrus, the planum temporale and the mid-to-posterior superior temporal gyrus. The results of a T-test on the positive effect of Familiarity (Familiar > Novel), a manipulation we hypothesized to reveal semantic processing, are presented in Table 3 and Figure 5, Row B. Greater activation for familiar

objects was observed within the occipito-temporal parietal junctions bilaterally, consistent with the results of the Familiar>Novel contrasts in the sound and visual form modalities separately (Figure 4, Rows A and B). In the left hemisphere, the cluster peaked in the posterior middle temporal gyrus where it abuts the extrastriate visual cortex. In the right hemisphere, the cluster peaked in the posterior superior temporal gyrus and extended more dorsally into the ventral parietal lobe than the left hemisphere cluster. The right lateralized cluster also extended rostrally into auditory cortex in the mid-to-posterior superior temporal gyrus. In the left hemisphere, auditory cortex was revealed in separate cluster in the planum temporale that also extended to the posterior insula. A right hemisphere cluster was observed in the parahippocampal gyrus, and area associated with high level visual processing and object recognition. Indeed, some of the regions showing an effect of Familiarity were located just anterior to those showing a main effect of modality suggesting that conceptual-level processing occurs upstream in the modality-specific pathways. In the superior temporal gyrus, the effect was observed in the most anterior aspect of the area showing the effect of modality. Familiarity also drove greater activation in lateral occipito-temporal regions adjacent to and including extras-striate regions around the posterior middle temporal gyrus, whereas effects of modality were in early visual cortex and medial occipito-temporal regions.

--Insert Figure 5 and Tables 2 and 3 about here---

Consistent with a role for the region in transmodal semantic processing, there was a cluster in both in the left and right anterior temporal lobe, peaking at the middle temporal gyrus, which responded to the familiarity manipulation (but did not reveal a modality effect). When each modality was considered separately, we only observed right ATL activation and for the visual form condition only. It is possible that the bilateral ATL activation in the present analysis reflects a greater statistical power to

detect subtle effects owing to collapsing across the conditions. In particular, the absence of ATL activation in the sound condition may reflect lower statistical power than the visual form condition due to a lower number of subjects. Moreover, as observed in the above analyses (Section 3.1), there was also greater activation in the medial prefrontal cortex, the precuneus and the posterior cingulate cortex. While it is possible that the results of this analysis reflect common processing across the presentation modalities, they could be predominantly driven by strong activation in only one of the two conditions (Nichols, Brett, Andersson, Wager, & Poline, 2005). Therefore, a conjunction analysis that reveals only voxels that were activated independently in both conditions, was necessary. The results of the conjunction analysis are displayed in Figure 5, Row C. It revealed the largely the same clusters (albeit much smaller, due to the much more stringent nature of the analysis), with the exception of the posterior cingulate cortex, which was absent. Although not clearly shown in Figure 5, there was bilateral posterior parahippocampal gyrus activation. The left ATL cluster was greatly reduced in size to two small separate clusters in the anterior MTG. This is likely driven by the lack of supra-threshold activation in the auditory modality, which again may reflect reduced power due to there being fewer subjects than in the visual form modality.

No interactions between Familiarity and Modality were observed in the whole-brain analysis. A strong form of our hypothesis regarding contributions of modality-specific regions to conceptual processing would be that an interaction should be observed such that they will respond equally to novel and familiar concepts in the associated input modality, whereas other presentation modalities will evoke activation only for familiar concepts that are more robustly represented across ‘hubs’ and ‘spoke’ regions. Statistically speaking, we believe this is unlikely to be observed, especially in a conservative whole-brain approach, as the effects of familiarity are in all probability much more subtle than modality effects and this has the potential to cloud an interaction effect. Additionally, the stimuli in each

presentation modality differ not only in modality but in perceptual richness and task difficulty. These confounds are likely to introduce noise that make direct comparisons between modalities difficult to interpret and interactions unlikely. As such, we focused primarily upon regional effects of familiarity independently within each modality.

### 3.3. Targeted regions of interest (ROI) analyses

The location of ROIs and the results of this analysis are displayed in Figure 6. The results are largely consistent with those in the whole-brain analyses above (Sections 3.1 and 3.2). Both the left and the right ATL ROIs were significantly more active for familiar entities regardless of whether the stimulus was the visual form of or the sound made by objects. In the whole-brain analyses in Section 3.1, only the right ATL was activated and only in the visual form condition. The greater statistical power afforded by the ROI analyses therefore appeared to greatly improve sensitivity to effects of familiarity.

--Insert Figure 6 about here--

Likewise, the superior temporal gyrus (STG) ROI, positioned over auditory association cortex, was significantly more active for familiar > novel entities when probed in the auditory domain and when probed by the visual form, albeit the effect was smaller in the latter. This pattern held for both the left and right hemisphere homologues. On the contrary, no significant effects of familiarity were observed in the ventral occipito-temporal cortex (vOTC) ROI, although there was a trend in the sound condition in the right hemisphere ROI ( $p = 0.11$ ). This ROI was situated more posteriorly than the activation observed in the whole-brain analysis and thus it may have not been optimally situated to detect such an effect. The whole-brain results implicated the mid parahippocampal gyri and the lateral extrastriate

visual cortex, further downstream in the visual processing hierarchy. This is consistent with prior observations that increased cortical activity of visual regions associated with recognition shifts more anteriorly with increased certainty of an object's identity (Bar et al., 2001).

#### **4. Discussion**

In this fMRI study, we set out to explore the roles of unimodal sensorimotor association areas and transmodal temporal association cortices in semantic processing. Distributed-only “embodied” theories propose that conceptualization is underpinned by reactivation of perception-action traces distributed across a network of remote unimodal association areas (Allport, 1985; Pulvermüller, Moseley, Egorova, Shebani, & Boulenger, 2014). An alternative theory suggests that in addition to the contribution of sensorimotor regions, or “spokes”, there is convergence of multi-modal information and further abstraction/computation within a centralized transmodal representational “hub”. Critically, the ‘hub-and-spoke’ model, as instantiated computationally by Rogers et al. (2004), requires the conjoint simultaneous action/computations of all the distributed ‘spokes’ and the ‘hub’ for conceptual representation. The hub is purported to be located specifically in the anterior temporal lobes, bilaterally. We tested a hypothesis congruent with the hub and spoke model, specifically that activation during semantic processing tasks should be observed both in a distributed set of unimodal association areas and the ATL. Moreover, this distributed pattern of activation should be observed irrespective of the input modality such that association areas not directly engaged by the stimulus should be activated in addition to those directly stimulated (e.g., auditory areas should be engaged by visual semantic stimuli). In order to disentangle activation related to semantic processing from that related to modality-specific perceptual processes, we employed a semantic manipulation of familiarity. Familiar and novel entities were closely matched in terms of perceptual complexity and therefore we hypothesized that any difference in

activation between these groups of entities would reflect ‘depth’ of semantic encoding rather than pre-semantic perceptual differences. Familiar object concepts are likely to have more coherent representations in the sensorimotor spokes, and as such recognition will evoke greater, more robust activations (see also Bar et al., 2001).

We interpret our findings as grossly supportive of the overarching predictions of the hub-and-spoke model. The bilateral ATL exhibited a greater response to familiar items compared to newly learned entities. The ROI analyses demonstrated that this ATL activation occurred irrespective of the input modality (environmental sound or visual form), a finding that is consistent with previous investigations delineating this cortical region as a transmodal hub (Vandenberghe et al., 1996; Visser, Jefferies, & Lambon Ralph, 2010). Our major finding relates to the contribution of unimodal spoke regions to semantic processing. We demonstrated that activation of auditory and visual association regions during a semantic task does not demand direct sensory stimulation. Auditory association cortex exhibited activation not only when semantic knowledge was probed by auditory stimuli but also, albeit less strongly, when it was probed via visual form. This result suggests that presentation of the visual form of, for example, a dog activates representations of its characteristic environmental sound. Our results regarding the contribution of visual association regions were less straightforward, although we believe they are not at odds with our hypothesis.

Our a priori hypothesis regarding the location of a visual spoke was derived from previous functional imaging studies that demonstrated semantic category-related patterns of activation in the ventral occipito-temporal cortex including the posterior fusiform gyrus (Aguirre, Zarahn, & D’esposito, 1998; Chao et al., 1999; Kanwisher, McDermott, & Chun, 1997). This region, situated anterior to early visual cortex (and the effect of modality in the present study), is associated with high-level visual processing, including object identification. Under the distributed-only framework, these category-related

effects have been interpreted as indicative of an exclusive role in visual semantic feature information, which renders it a candidate visual spoke region (e.g., Martin et al., 2000). However, we failed to observe any significant differential response to familiar objects compared to newly learned objects within this region of interest that would support a conceptual-level function, particularly one of representing object-specific information. Category-related patterns of activity may, therefore, reflect a function that is limited to global visual feature conjunctions shared across many category-exemplars. Tyler and colleagues (2004) have argued that there exists a gradient of specificity along the posterior-anterior axis of the ventral temporal lobe, such that posterior regions encode gross domain distinctions with progressive conjunctions of features occurring along the anterior axis, giving rise to specific person and object-knowledge distinctions in the most anterior portions of the ventral visual pathway, including perirhinal cortex and proximal regions such as parahippocampal cortex (see also Bar et al., 2001). Our whole-brain analyses revealed greater activation for familiar objects in the medial temporal cortex bilaterally. This raises the possibility that the visual spoke is located within a more anterior distribution to our ROI. The effect of familiarity in the medial temporal lobe was, however, only observed in the visual form condition. Therefore, our hypothesis that an object's sound would activate visual association regions reflecting recapitulation of its form was not met in this region. Still, failure to reject the null hypotheses may stem from the lower number of subjects in the sound condition and a lack of statistical power to detect this effect in this region.

The lateral posterior temporal and ventral parietal cortex abutting the lateral extrastriate visual cortex exhibited an effect of familiarity in both the visual form and sound conditions. The lack of a modality-specific effect implicates some of this region as omni-modal in nature. Indeed, this area has been described as heteromodal cortex important for conceptual representation (Bonner, Peelle, Cook, & Grossman, 2013; Price, Bonner, Peelle, & Grossman, 2015) and the fact that it is located between dorsal

auditory region and ventral visual regions make it ideally suited for cross-modal integration (Beauchamp, Argall, Bodurka, Duyn, & Martin, 2004; Binney, Parker, & Lambon Ralph, 2012). It has also been implicated as part of a network subserving an executive component of semantic processing (Jefferies, 2013; Noonan, Jefferies, Visser, & Lambon Ralph, 2013). On the other hand, this region has also been demonstrated to exhibit category-specific patterns of activation (Chao et al., 1999) and thus might also be considered a candidate spoke region, possibly even a visual spoke. Specifically though, it has been suggested that it may be site for stored information about object motion (Martin, 2000; Beauchamp et al, 2003). Activation in this region could reflect retrieval of object motion features. However, the present results could feasibly be explained by either one of these accounts, and our study design cannot adjudicate between either.

Analysis of the point-light motion condition failed to reveal any significant effects of familiarity. This brings into question whether the stimuli were suited to detect neural responses associated with access to object-specific information. Point light displays have been successfully demonstrated to evoke responses in the lateral posterior temporal cortex (Beauchamp, Lee, Haxby, & Martin, 2003) but these responses were modulated by category-level distinctions (biological vs. non-biological motion) that we did not explicitly contrast in the present study. Moreover, the lateral posterior temporal cortex may be rather less tuned to finer-grained within-category distinctions probed by familiarity. It is unclear, however, why there was no clear response to familiar and newly learned concepts in the whole brain analysis (with the exception of a small motor cortex cluster). One possibility is that point light may reflect an ecologically unnatural means to probe object recognition and subsequent conceptualization compared to the form and sound conditions (i.e., we have no evolutionary precedent for naming objects depicted from point light). As such, point light stimuli may elicit high inter-subject and inter-item variability that consequently produce null results at the group level.

The cognitive subtraction of novel from familiar objects offers a unique approach to isolating semantic activation. We observed such activation in regions considered to be some of the key components of the core network subserving semantic cognition and language comprehension (Jefferies, 2013; Turken & Dronkers, 2011), including the bilateral ATL, the posterior middle temporal gyrus and the temporo-parietal junction, which supports the effectiveness of this approach. This also suggests that greater activation for familiar relative to novel objects observed in unimodal sensory association regions reflects semantic upregulation of these regions. It remains to be demonstrated, however, whether this reflects processes key to conceptualization or other cognitive or perceptual processes that are differentially engaged by familiar objects/concepts but not directly associated with conceptual processing (see also Mahon, 2014; Reilly et al., 2014). Limitations in the temporal resolution of fMRI preclude addressing these contingencies directly in the present study. In lieu of future investigations that will, some alternatives are worthy of consideration. Rather than reflecting a process of conceptualization, greater activation of a directly stimulated modality-specific region by familiar compared to newly learned objects could reflect greater semantic feedback to sensorimotor regions acting as a top-down influence on high-level perceptual processing (Hon, Thompson, Sigala, & Duncan, 2009). When the modality-specific region is not directly engaged by the stimulus, the differential activity could be interpreted as reflecting enhanced mental imagery or perceptual simulation (although not necessarily at the level of conscious awareness). This might follow from familiar items having more coherent representation in the sensorimotor domain and richer associations in episodic memory. Indeed, we also observed greater activation for familiar objects in posterior ventral cingulate cortex which has been linked to episodic memory processes, including the internal generation and maintenance of mental imagery (Vann, Aggleton, & Maguire, 2009). This is to say that activation of action-perception traces and episodic associations could be a result of feedforward propagation and secondary to conceptual

processing, playing an elaborative rather than integral role. Mahon (2014) likened this phenomenon to the domain of speech perception where a common finding is that as we hear words, their corresponding orthographic representations are rapidly activated. This form of ‘passive priming’ may occur, but few would argue that activation of orthography is a necessary pre- or co-requisite for effective speech perception.

Contrasting familiar and novel entities to reveal semantic activation may also come with pitfalls. In particular, naming familiar entities is likely to be easier than naming novel entities and thus contrasting the two has the potential to yield activation patterns not entirely associated with conceptual processing. Such contrasts may invoke regions considered to play a functional role in the Default Mode Network (DMN), for example. The DMN is an anatomically-defined network that shows deactivation during goal-directed tasks as compared to rest, with some component regions showing stronger task-induced deactivations as task difficulty increases (Humphreys, Hoffman, Visser, Binney, & Lambon Ralph, 2015; Seghier & Price, 2012). The DMN includes medial prefrontal cortex, the precuneus, the posterior cingulate cortex, the hippocampus and the temporo-parietal junction (including the angular gyrus) (Buckner et al., 2008). Many of these DMN regions were activated in the familiar > novel contrasts we reported here. One explanation for this is that during rest, the brain is afforded more opportunity for task-unrelated thought, which may engage processes such as internal speech and mental imagery (Binder et al., 1999). Along the same lines, if the familiar items are easier to name, then component of the DMN may appear to activate more during these trials than in novel item trials not because they are more engaged but because they are less engaged in the task. Although we trained participants on both familiar and novel items to 100% accuracy, it remains possible/probable that naming novel items was more cognitively demanding than naming familiar items. Thus, task difficulty

poses an alternative explanation for the large areas of activation we observed in the posterior cingulate cortex, precuneus and medial frontal cortex.

In conclusion, we provide novel empirical support for the hypothesis that both the ATL hub and modality-specific spokes are engaged conceptual processing. Future work will address some of the above outlined issues and determine whether the observed involvement of modality-specific ‘spoke’ regions is *necessary* for performance of semantic tasks. The question of necessity arguably cannot be answered through fMRI or any other single experimental method currently at our disposal. For example, in the current study, we cannot know whether spokes are engaged in parallel or are instead sequentially “turned on” by rapid distillation through a hub. The question of necessity for spoke activation will be gleaned through converging and perhaps tandem evidence from a variety of sources with varying scales of temporal and spatial resolution (e.g., MEG, TMS, neuropsychology, fMRI).

### **ACKNOWLEDGEMENTS**

We thank Alison O'Donoghue for her assistance in scheduling and recruiting participants. This study was supported by US Public Health Service Grants R01 DC013063 (JR), and T32 T32AG020499 (AG).

## REFERENCES

- Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to “building” stimuli: evidence and implications. *Neuron*, *21*(2), 373–383.
- Allport, D. A. (1985). Distributed memory, modular subsystems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia* (pp. 207–244). Edinburgh: Churchill Livingstone.
- Bar, M., Tootell, R. B., Schacter, D. L., Greve, D. N., Fischl, B., Mendola, J. D., ... Dale, A. M. (2001). Cortical mechanisms specific to explicit visual object recognition. *Neuron*, *29*(2), 529–535.
- Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004). Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. *Nature Neuroscience*, *7*(11), 1190–1192.
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). fMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, *15*(7), 991–1001. <http://doi.org/10.1162/089892903770007380>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where Is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: a functional MRI study. *Journal of Cognitive Neuroscience*, *11*(1), 80–93.
- Binney, R. J., Embleton, K. V., Jefferies, E., Parker, G. J. M., & Lambon Ralph, M. A. (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. *Cerebral Cortex*, 20(11), 2728–2738. <http://doi.org/10.1093/cercor/bhq019>

- Binney, R. J., Parker, G. J. M., & Lambon Ralph, M. A. (2012). Convergent connectivity and graded specialization in the rostral human temporal lobe as revealed by diffusion-weighted imaging probabilistic tractography. *Journal of Cognitive Neuroscience*, 24(10), 1998–2014.
- Bonner, M. F., Peelle, J. E., Cook, P. A., & Grossman, M. (2013). Heteromodal conceptual processing in the angular gyrus. *Neuroimage*, 71, 175–186.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, 38, 1207–1215.
- Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage*, 16(2), S497.
- Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network. *Annals of the New York Academy of Sciences*, 1124(1), 1–38.
- Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, 2(10), 913–919.
- Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, 12, 478–484.
- Chao, L. L., Weisberg, J., & Martin, A. (2002). Experience-dependent modulation of category-related cortical activity. *Cerebral Cortex*, 12(5), 545–51.
- Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Lambon Ralph, M. A. (2004). Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia. *Cognitive Neuropsychology*, 21(5), 513–527.

- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*(1-2), 25–62.
- Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., ... Tyler, L. K. (2000). Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. *Neuroimage*, *11*(6), 589–600.
- Eggert, G. H. (1977). *Wernicke's Works on Aphasia: A Sourcebook and Review* (Vol. 1). The Hague, Netherlands: Mouton.
- Embleton, K. V., Haroon, H. A., Morris, D. M., Ralph, M. A. L., & Parker, G. J. (2010). Distortion correction for diffusion-weighted MRI tractography and fMRI in the temporal lobes. *Human Brain Mapping*, *31*(10), 1570–1587.
- Gage, N., & Hickok, G. (2005). Multiregional cell assemblies, Temporal binding and the representation of conceptual knowledge in cortex: A modern theory by a “classical” neurologist, Carl Wernicke. *Cortex*, *41*(1998), 823–832.
- Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3), 455–479.  
<http://doi.org/10.1080/02643290442000310>
- Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *The Journal of Neuroscience*, *26*(18), 4917–4921.
- Goll, J. C., Crutch, S. J., Loo, J. H., Rohrer, J. D., Frost, C., Bamiou, D.-E., & Warren, J. D. (2010). Non-verbal sound processing in the primary progressive aphasia. *Brain*, *133*(1), 272–285.
- Halgren, E., Wang, C., Schomer, D. L., Knake, S., Marinkovic, K., Wu, J., & Ulbert, I. (2006). Processing stages underlying word recognition in the anteroventral temporal lobe. *Neuroimage*, *30*(4), 1401–1413.

- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 207–301.
- Hoenig, K., Sim, E. J., Bochev, V., Herrnberger, B., & Kiefer, M. (2008). Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience*, *20*(10), 1799–1814.
- Hoffman, P., Jones, R. W., & Ralph, M. A. (2012). The degraded concept representation system in semantic dementia: damage to pan-modal hub, then visual spoke. *Brain*, *135*(Pt 12), 3770–3780.  
<http://doi.org/10.1093/brain/aws282>
- Hon, N., Thompson, R., Sigala, N., & Duncan, J. (2009). Evidence for long-range feedback in target detection: Detection of semantic targets modulates activity in early visual areas. *Neuropsychologia*, *47*(7), 1721–1727.
- Humphreys, G. F., Hoffman, P., Visser, M., Binney, R. J., & Lambon Ralph, M. A. (2015). Establishing task-and modality-dependent dissociations between the semantic and default mode networks. *Proceedings of the National Academy of Sciences*, *112*(25), 7857–7862.
- Humphreys, G. F., & Lambon Ralph, M. A. (2014). Fusion and Fission of Cognitive Functions in the Human Parietal Cortex. *Cerebral Cortex*. <http://doi.org/10.1093/cercor/bhu198>
- Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex*, *49*(3), 611–625.  
<http://doi.org/10.1016/j.cortex.2012.10.008>
- Jouen, A. L., Ellmore, T. M., Madden, C. J., Pallier, C., Dominey, P. F., & Ventre-Dominey, J. (2014). Beyond the word and image: characteristics of a common meaning system for language and vision revealed by functional and structural imaging. *NeuroImage*.  
<http://doi.org/10.1016/j.neuroimage.2014.11.024>

- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *361*(1476), 2109–2128.
- Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cogn Affect Behav Neurosci*, *1*(3), 207–221.
- Kemmerer, D. (n.d.). Are the motor features of verb meanings represented in the precentral motor cortices? Yes, but within the context of a flexible, multilevel architecture for conceptual knowledge. *Psychonomic Bulletin and Review*.
- Kemmerer, D., Castillo, J. G., Talavage, T., Patterson, S., & Wiley, C. (2008). Neuroanatomical distribution of five semantic components of verbs: Evidence from fMRI. *Brain and Language*, *107*(1), 16–43.
- Lambon Ralph, M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philosophical Transactions of the Royal Society B*, *369*, 20120392.
- Lambon Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological evidence and a computational model. *Journal of Cognitive Neuroscience*, *13*(3), 341–356.
- Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and differentiation in semantic memory: insights from semantic dementia. *Annals of the New York Academy of Science*, *1124*, 61–76.  
<http://doi.org/10.1196/annals.1440.006>

- Lambon Ralph, M. A., Pobric, G., & Jefferies, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: convergent evidence from rTMS. *Cerebral Cortex*, *19*(4), 832–8. <http://doi.org/10.1093/cercor/bhn131>
- Lambon Ralph, M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences USA*, *107*(6), 2717–2722. <http://doi.org/10.1073/pnas.0907307107>
- Mahon, B. Z. (2014). What is embodied about cognition? *Language, Cognition, and Neuroscience*.
- Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, *38*(3), 487–497.
- Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., ... Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain*, *133*(11), 3256–3268. <http://doi.org/awq272> [pii] 10.1093/brain/awq272 [doi]
- Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, *25*(3), 653–660.
- Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (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*(11), 1824–1850. [http://doi.org/10.1162/jocn\\_a\\_00442](http://doi.org/10.1162/jocn_a_00442)
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987.

- Piwnica-Worms, K. E., Omar, R., Hailstone, J. C., & Warren, J. D. (2010). Flavour processing in semantic dementia. *Cortex*, *46*(6), 761–768.
- Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2007). Anterior temporal lobes mediate semantic representation: mimicking semantic dementia by using rTMS in normal participants. *Proceedings of the National Academy of Sciences, U S A*, *104*(50), 20137–20141.  
<http://doi.org/0707383104> [pii] 10.1073/pnas.0707383104 [doi]
- Pobric, G., Jefferies, E., & Lambon Ralph. (2010). Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Current Biology*, *20* (10), 964-968.
- Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging Evidence for the Neuroanatomic Basis of Combinatorial Semantics in the Angular Gyrus. *The Journal of Neuroscience*, *35*(7), 3276–3284. <http://doi.org/10.1523/JNEUROSCI.3446-14.2015>
- Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, *5*(12), 517–524. [http://doi.org/10.1016/s1364-6613\(00\)01803-9](http://doi.org/10.1016/s1364-6613(00)01803-9)
- Pulvermüller, F. (2013). Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. *Brain and Language*, *127*(1), 86–103.  
<http://doi.org/10.1016/j.bandl.2013.05.015>
- Pulvermüller, F., Moseley, R. L., Egorova, N., Shebani, Z., & Boulenger, V. (2014). Motor cognition-motor semantics: Action perception theory of cognition and communication. *Neuropsychologia*, *55*, 71–84. <http://doi.org/10.1016/j.neuropsychologia.2013.12.002>
- Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain Signatures of Meaning Access in Action Word Recognition. *Journal of Cognitive Neuroscience*, *17*(6), 884–892.
- Reilly, J., Harnish, S., Garcia, A., Hung, J., Rodriguez, A. D., & Crosson, B. (2014). Lesion symptom mapping of manipulable object naming in nonfluent aphasia: can a brain be both embodied and

disembodied? *Cognitive Neuropsychology*, 31(4), 287–312.

<http://doi.org/10.1080/02643294.2014.914022>

Reilly, J., Peelle, J. E., Garcia, A., & Crutch, S. J. (2016). Linking somatic and symbolic representation in semantic memory: The Dynamic Multilevel Reactivation framework. *Psychonomic Bulletin & Review*.

Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and Deterioration of Semantic Memory: A Neuropsychological and Computational Investigation. *Psychological Review*, 111(1), 205–235.

Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA US: MIT Press. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=psych&AN=2004-18753-000&site=ehost-live>

Seghier, M. L., & Price, C. J. (2012). Functional heterogeneity within the default network during semantic processing and speech production. *Front Psychol*, 3(10.3389). Retrieved from [https://www.researchgate.net/profile/Mohamed\\_Seghier2/publication/230699777\\_Functional\\_Heterogeneity\\_within\\_the\\_Default\\_Network\\_during\\_Semantic\\_Processing\\_and\\_Speech\\_Production/links/544a668b0cf21cdb24f773d6.pdf](https://www.researchgate.net/profile/Mohamed_Seghier2/publication/230699777_Functional_Heterogeneity_within_the_Default_Network_during_Semantic_Processing_and_Speech_Production/links/544a668b0cf21cdb24f773d6.pdf)

Shimotake, A., Matsumoto, R., Ueno, T., Kunieda, T., Saito, S., Hoffman, P., ... others. (2014). Direct exploration of the role of the ventral anterior temporal lobe in semantic memory: cortical stimulation and local field potential evidence from subdural grid electrodes. *Cerebral Cortex*, bhu262.

- Shtyrov, Y., & Pulvermüller, F. (2007). Early MEG activation dynamics in the left temporal and inferior frontal cortex reflect semantic context integration. *Journal of Cognitive Neuroscience*, *19*(10), 1633–1642.
- Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex*, *15*(10), 1602–1608. <http://doi.org/bhi038> [pii] 10.1093/cercor/bhi038 [doi]
- Taylor, K. I., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2006). Binding crossmodal object features in perirhinal cortex. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *103*(21), 8239–8244.
- Thierry, G., & Price, C. J. (2006). Dissociating verbal and nonverbal conceptual processing in the human brain. *Journal of Cognitive Neuroscience*, *18*(6), 1018–1028.
- Tranel, D., Grabowski, T. J., Lyon, J., & Damasio, H. (2005). Naming the same entities from visual or from auditory stimulation engages similar regions of left inferotemporal cortices. *Journal of Cognitive Neuroscience*, *17*(8), 1293–1305. <http://doi.org/10.1162/0898929055002508> [doi]
- Trumpp, N. M., Kliese, D., Hoenig, K., Haarmeier, T., & Kiefer, M. (2012). Losing the sound of concepts: Damage to auditory association cortex impairs the processing of sound-related concepts. *Cortex*. [http://doi.org/S0010-9452\(12\)00057-3](http://doi.org/S0010-9452(12)00057-3) [pii] 10.1016/j.cortex.2012.02.002 [doi]
- Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in Systems Neuroscience*, *5*(February), 1. <http://doi.org/10.3389/fnsys.2011.00001>
- Tyler, L. K., Chiu, S., Zhuang, J., Randall, B., Devereux, B. J., Wright, P., ... Taylor, K. I. (2013). Objects and categories: feature statistics and object processing in the ventral stream. *Journal of Cognitive Neuroscience*, *25*(10), 1723–1735.

- Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*(6597), 254–6.
- Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, *10*(11), 792–802.
- Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, *22*(6), 1083–1094.
- Visser, M., & Lambon Ralph, M. A. (2011). Differential contributions of bilateral ventral anterior temporal lobe and left anterior superior temporal gyrus to semantic processes. *Journal of Cognitive Neuroscience*, *23*(10), 3121–3131. [http://doi.org/10.1162/jocn\\_a\\_00007](http://doi.org/10.1162/jocn_a_00007)
- Warren, J. D., Jennings, A. R., & Griffiths, T. D. (2005). Analysis of the spectral envelope of sounds by the human brain. *Neuroimage*, *24*(4), 1052–1057.
- Wernicke, C. (1874). *Der aphasische symptomemkomplex: Eine psychologische Studie auf anatomischer basis*. Breslau: Cohn und Weigert.
- Wheatley, T., Weisberg, J., Beauchamp, M. S., & Martin, A. (2005). Automatic priming of semantically related words reduces activity in the fusiform gyrus. *Journal of Cognitive Neuroscience*, *17*(12), 1871–1885.
- Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences USA*, *104*(15), 6430–6435. <http://doi.org/0607061104> [pii] 10.1073/pnas.0607061104 [doi]

**Table 1. Whole-brain analyses of greater activation for familiar concepts relative to novel concepts in each presentation modality**

Modality	Brain Region		Cluster Size	Peak Z Value	Peak MNI Coordinates		
					X	Y	Z
<b>Form</b>							
L	vOTC and medial temporal lobe	**	160				
	posterior parahippocampal cortex			4.002	-27	-40	-20
	mid parahippocampal cortex			3.658	-30	-31	-23
	anterior hippocampus			3.472	-15	-13	-23
L	Occipital-temporal-parietal junction	**	147				
	superior posterior temporal cortex			4.139	-45	-43	1
	lateral occipito-temporal cortex (MTG)			3.906	-48	-61	16
	extrastriate visual cortex			3.497	-42	-79	19
L	Orbitofrontal cortex		64				
	rostral OFC			4.301	-24	38	-11
	caudal OFC			4.017	-21	26	-11
L	Planum temporale and insula		50				
	posterior insula			3.753	-36	-16	-2
	medial temporal			2.632	-27	-16	-8
L/R	Frontal, temporal and parietal lobes	**	2968				
	orbitofrontal cortex			5.180	24	32	-14
	mid parahippocampal cortex			4.919	30	-25	-20
	anterior MTG			4.847	54	-1	-26
R	Thalamus		70				
	caudal thalamus			3.820	6	-22	4
	ventrolateral thalamus			3.594	15	-13	4
	caudal thalamus			3.347	18	-28	7
<b>Sound</b>							
L	Occipital-temporal-parietal junction	**	238				
	extrastriate visual cortex			4.800	-51	-70	4
	lateral occipito-temporal cortex (MTG)			4.211	-45	-61	10
	supramarginal gyrus/posterior STG			3.956	-60	-49	25
L	Planum temporale and insula		98				
	posterior insula			3.648	-48	-7	7
	insula/medial STG			3.436	-39	-19	-2
	frontal operculum			2.615	-60	-4	13
L	Posterior STG		32				
				3.311	-54	-31	13
L/R	Bilateral medial PFC	**	480				
	left medial PFC			3.950	-6	62	10
	right medial PFC			3.776	12	65	10

		right medial PFC			3.738	6	59	19
L/R		Bilateral precuneus/posterior cingulate	**	212				
		right precuneus			5.376	6	-52	37
		left precuneus			4.380	-9	-55	40
L/R		Bilateral retrosplenial cortex	*	135				
		retrosplenial cortex			3.745	0	-58	10
		right retrosplenial cortex			3.507	21	-61	16
		left retrosplenial cortex			2.665	-18	-58	7
R		Occipital-temporal-parietal junction	**	561				
		supramarginal gyrus/posterior STG			3.894	60	-49	22
		posterior superior temporal sulcus			3.845	45	-61	13
		anterodorsal supramarginal gyrus			3.640	60	-28	22
R		Inferolateral occipito temporal cortex		46				
		lateral occipito-temporal cortex			3.855	60	-49	-8
		ventral occipito-temporal cortex			3.179	54	-46	-14
		lateral occipital cortex			2.790	51	-55	-11
Motion								
L		Left ventral central sulcus		44				
					3.184	-45	-22	34
					2.793	-36	-19	34
R		Right ventral postcentral gyrus		89				
					3.460	48	-10	34
					3.032	36	-13	34
					2.993	60	-10	28

Table shows up to 3 local maxima per cluster more than 8.0mm apart. All areas surviving  $p=0.005$ , uncorrected voxel-level threshold (minimum cluster size = 30 voxels) in each whole-brain analysis are reported. \*\* Areas surviving FDR-corrected cluster threshold at  $p<0.05$ . \* Areas surviving FDR-corrected cluster threshold at  $p<0.08$ . L = left; R = right; OFC = orbitofrontal cortex; PFC = prefrontal cortex; MTG = middle temporal gyrus; STG = superior temporal gyrus.

**Table 2. Whole-brain analyses of Main effect of Modality (Sound versus Form only)**

Brain Region	Cluster Size	Peak F Value	Peak Z Value	Peak MNI Coordinates		
				X	Y	Z
Left superior temporal cortex	465					
Mid planum temporale		56.62	6.15	-54	-19	1
Posterior superior temporal gyrus		39.90	5.36	-54	-34	10
Mid superior temporal gyrus		37.56	5.23	-57	-25	7
Bilateral Occipito-temporal cortex	3087					
left ventral occipito-temporal cortex (FG)		86.00	7.14	-30	-61	-17
right ventral occipital cortex (FG)		76.83	6.87	27	-76	-14
right ventral occipito-temporal cortex (FG)		73.84	6.77	33	-58	-17
Right superior temporal cortex	300					
Mid STG/planum temporale		54.94	6.08	60	-19	4

Clusters reported all survived a  $p < 0.001$ , uncorrected voxel-level threshold and a  $P < 0.05$  FDR-corrected cluster extent threshold. Up to 3 local maxima per cluster more than 8.0mm apart are reported. STG = superior temporal gyrus; FG = fusiform gyrus.

**Table 3. Whole-brain analyses of positive main effect of Familiarity (Familiar > Novel) across the Form and Sound presentation modalities**

Brain Region	Cluster Size	Peak T Value	Peak Z Value	Peak MNI Coordinates		
				X	Y	Z
L Occipital-temporal-parietal junction	211					
lateral occipito-temporal cortex (MTG)		6.08	5.33	-51	-67	22
supramarginal gyrus/posterior STG		4.15	3.87	-60	-52	25
L Planum temporale/posterior insula	129					
planum temporale		4.66	4.27	-45	-13	-2
L Anterior temporal cortex	68					
anterior MTG		4.86	4.44	-57	-1	-23
dorsal medial temporal pole		4.59	4.22	-36	17	-26
anterior MTG		4.38	4.05	-51	5	-32
L/R Bilateral medial PFC	986					
right ventromedial PFC		6.42	5.56	6	50	-5
left medial PFC		6.29	5.47	-3	65	7
bilateral anterior cingulate cortex		5.89	5.19	0	41	1
L/R Bilateral retrosplenial cortex	291					
right retrosplenial cortex		5.79	5.12	6	-55	13
left retrosplenial cortex		5.19	4.68	-9	-52	7
L/R Bilateral precuneus/posterior cingulate	138					
left precuneus		4.80	4.39	-6	-55	37
right posterior cingulate		3.69	3.48	6	-40	37
right precuneus		3.54	3.36	9	-49	40
R Occipital-temporal-parietal junction	885					
lateral occipito-temporal cortex (STG)		7.57	6.28	57	-61	19
posterior superior temporal sulcus		7.51	6.25	48	-61	16
posterior superior temporal sulcus		6.83	5.83	48	-49	10
R Anterior temporal cortex	138					
lateral temporal pole/anterior MTG		6.02	5.29	48	11	-29
anterior MTG		5.92	5.21	57	-1	-23
R Parahippocampal gyrus	98					
posterior parahippocampal gyrus		4.92	4.48	27	-37	-14
mid parahippocampal gyrus		4.30	3.99	27	-19	-20

Clusters reported all survived a  $p < 0.001$ , uncorrected voxel-level threshold and a  $P < 0.05$  FDR-corrected cluster extent threshold. Up to 3 local maxima per cluster more than 8.0mm apart are reported. L = left. R = right. PFC = prefrontal cortex; STG = superior temporal gyrus; MTG = middle temporal gyrus.

**Figure Captions**

Figure 1: Illustration of two competing models of cortical networks underpinning semantic memory. See main text for further description.

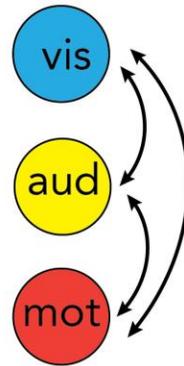
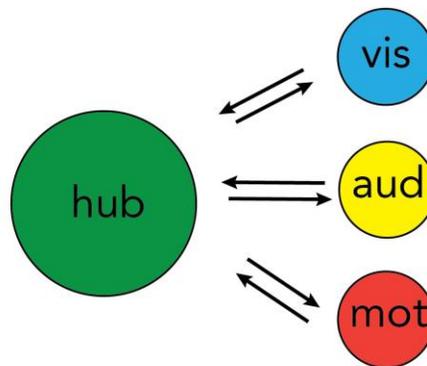
**1. Fully Distributed****2. Distributed + Hub**

Figure 2: Stimuli characteristics

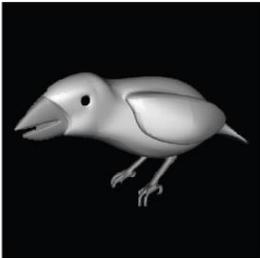
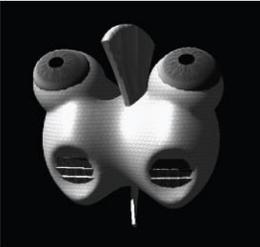
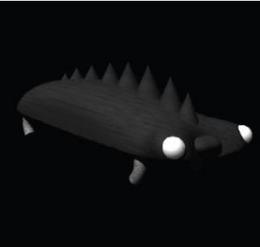
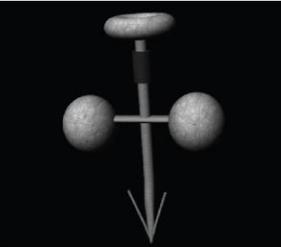
	bird	dog	clock	scissors	
familiar visual					
	plufky	korbok	blerga	wilzig	
novel visual					
motion	path	left-right	back-front	down-left	up-down
	manner	sweep-flip	sidewinding	spinning handle	drop-rotate
sound	High-freq Scream	Low-freq Growl	Whoosh	Ratcheting + Wooden Impact	

Figure 3: Example fMRI trial procedure

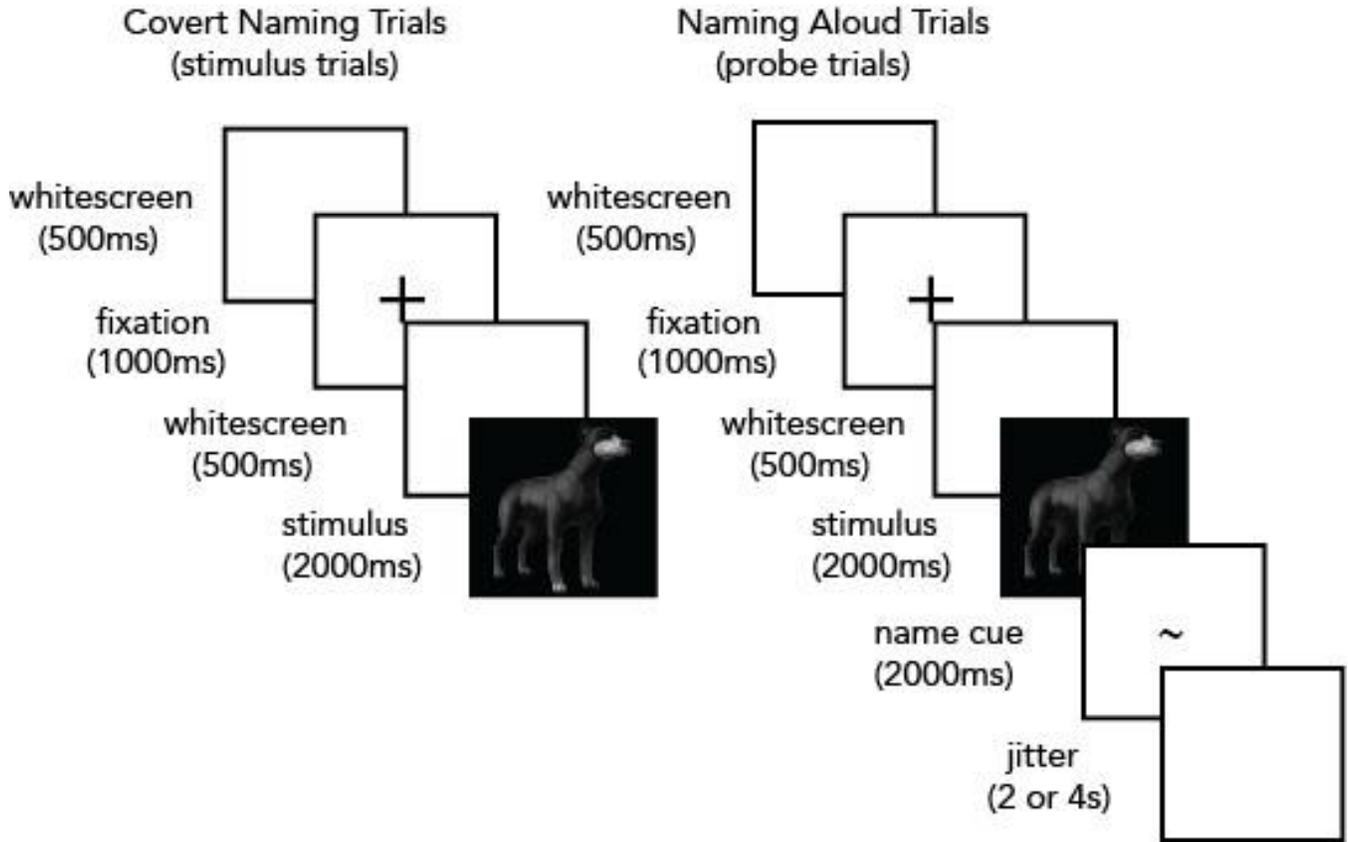


Figure 4: Whole-brain analyses of familiar minus novel concept processing probed by (a) environmental sound, (b) visual form, and (c) point-light motion. Statistical maps are displayed on a glass brain rendering with an uncorrected voxel-height threshold of  $p < 0.005$  and a minimum cluster extent of 30 voxels. Overlay brightness indicates distance from cortical surface.

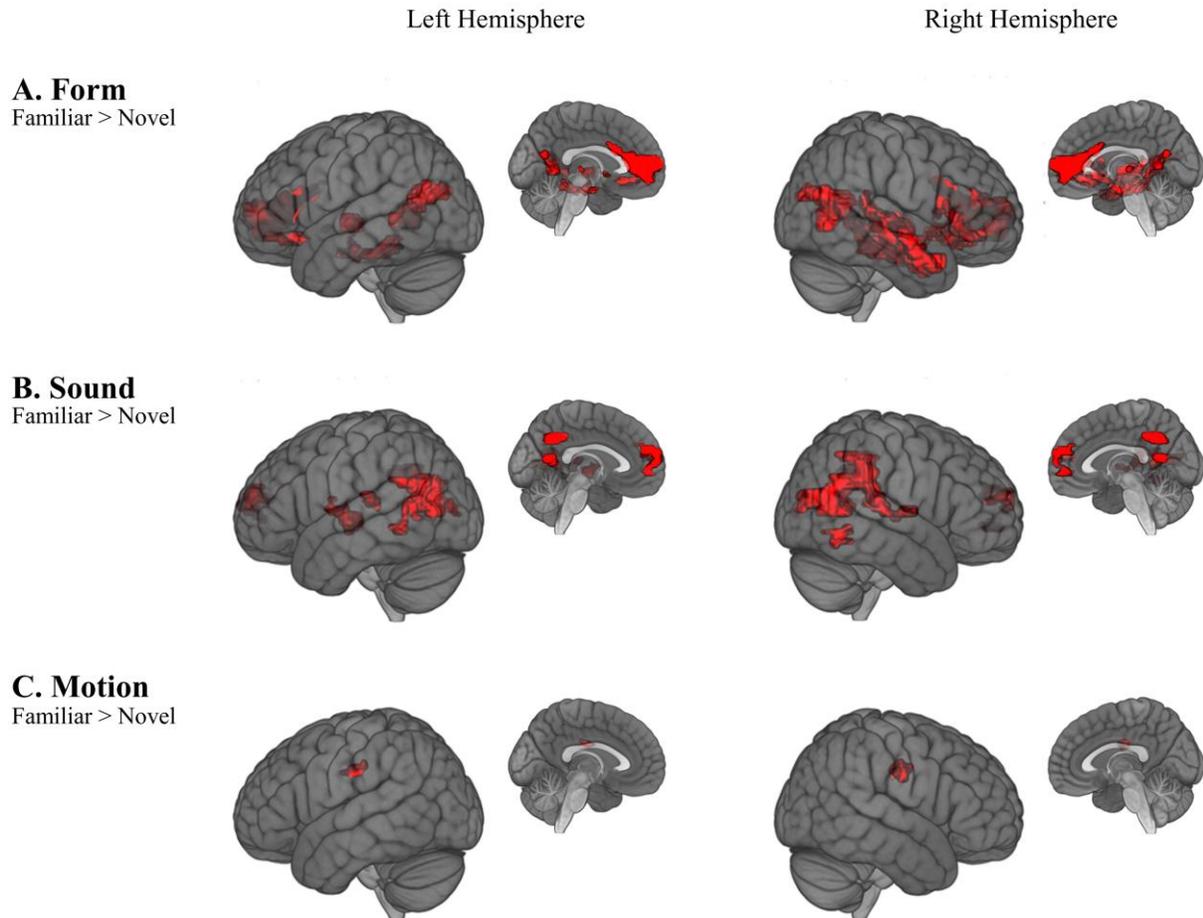


Figure 5: Whole-brain analysis of common activation to familiar > novel entities across the visual form and sound conditions. (a) F-test on the main effect of Modality, (b) T-test on the positive main effect of familiarity. (c) Conjunction of the independent analyses of the familiar > novel contrast in the sound and form conditions. The statistical maps in rows (a) and (b) are thresholded at an uncorrected voxel-height threshold of  $P < 0.001$  and a minimum cluster extent at  $p < 0.05$ , FDR-corrected. The conjunction was performed following application of an uncorrected threshold of  $p < 0.01$  to each of the two statistical maps in Figure 4, panels a and b. See text for further details. Overlay brightness indicates distance from cortical surface.

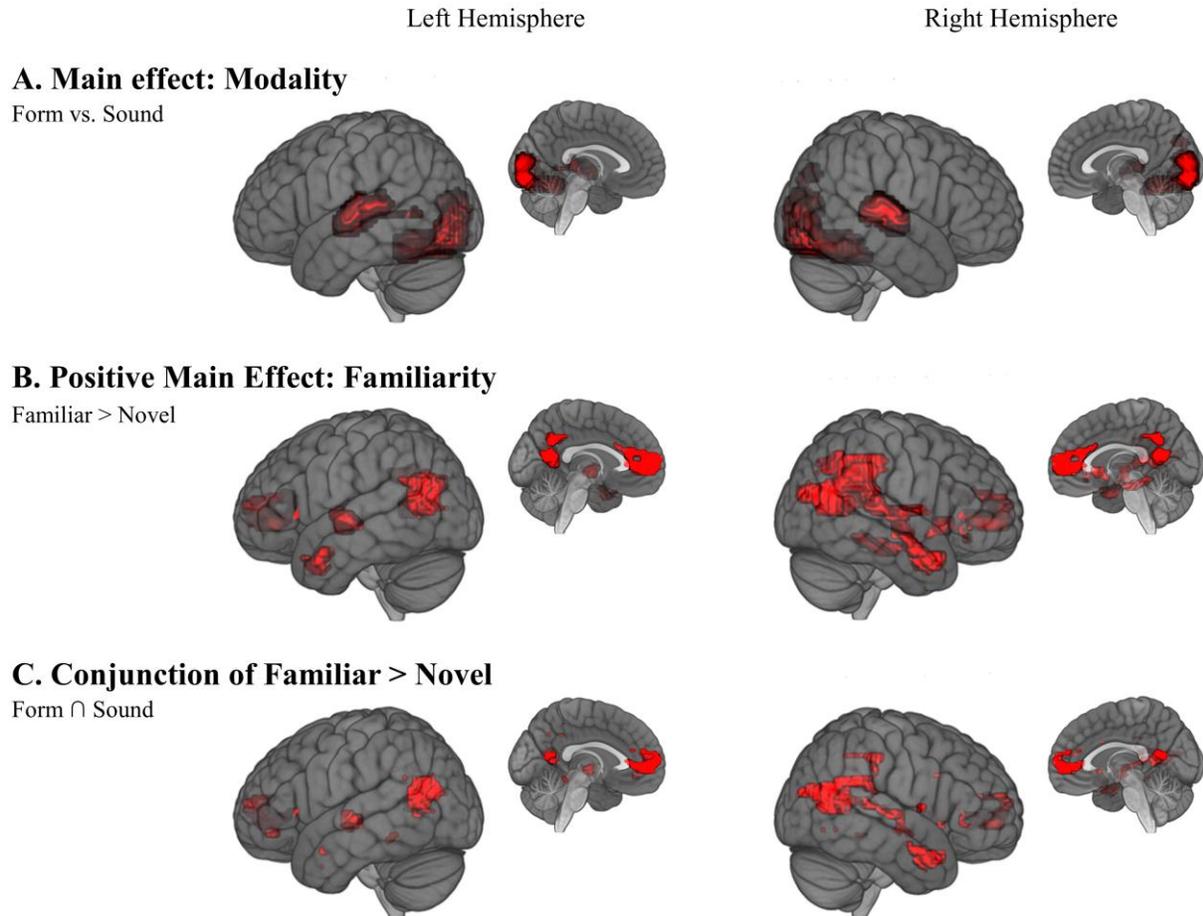


Figure 6: Contrast estimates (and standard measurement error) for familiar minus novel concepts in the auditory and visual presentation modality are displayed for a priori targeted regions of interest (ROIs) in each hemisphere. Significant positive values indicate greater activation for familiar concepts, significant negative values indicate greater activation for novel concepts, while non-significant values suggest no differential activation. See text for further details. \*  $p < 0.05$ .  $\sim p < 0.11$ . STG = superior temporal gyrus. ATL = anterior temporal lobe. vOTC = ventral occipito-temporal cortex.

