This is a repository copy of *Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/110309/

Version: Published Version

**Article:**
Murphy, Charlotte, Rueschemeyer, Shirley Ann orcid.org/0000-0001-6432-0917, Watson, David et al. (3 more authors) (2017) *Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality*. Neuroimage. pp. 19-31. ISSN 1053-8119

https://doi.org/10.1016/j.neuroimage.2016.11.067

**Reuse**
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

**Takedown**
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.
Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality

Charlotte Murphy*, Shirley-Ann Rueschemeyer, David Watson, Theodoros Karapanagiotidis, Jonathan Smallwood, Elizabeth Jefferies

Department of Psychology and York Neuroimaging Centre, University of York, UK

**Abstract**

Words activate cortical regions in accordance with their modality of presentation (i.e., written vs. spoken), yet there is a long-standing debate about whether patterns of activity in any specific brain region capture modality-invariant conceptual information. Deficits in patients with semantic dementia highlight the anterior temporal lobe (ATL) as an amodal store of semantic knowledge but these studies do not permit precise localisation of this function. The current investigation used multiple imaging methods in healthy participants to examine functional dissociations within ATL. Multi-voxel pattern analysis identified spatially segregated regions: a response to input modality in anterior superior temporal gyrus (aSTG) and a response to meaning in more ventral anterior temporal lobe (vATL). This functional dissociation was supported by resting-state connectivity that found greater coupling for aSTG with primary auditory cortex and vATL with the default mode network. A meta-analytic decoding of these connectivity patterns implicated aSTG in processes closely tied to auditory processing (such as phonology and language) and vATL in meaning-based tasks (such as comprehension or social cognition). Thus we provide converging evidence for the segregation of meaning and input modality in the ATL.

**Keywords:** Anterior temporal lobe (ATL), Multi-voxel pattern analysis (MVPA), Semantic Hub, Spoke, Resting-state connectivity

**1. Introduction**

Current neurocognitive models propose that concepts are represented in a large-scale distributed network comprising (1) sensory and motor ‘spoke’ regions that store knowledge of physical features and (2) convergence zones that integrate across multiple modalities (e.g., visual vs. auditory) to form abstract amodal representations (Damasio, 1989; Patterson et al., 2007). For example, the hub and spoke model of Patterson et al. (2007) proposes that information from modality-specific spoke regions is integrated in an amodal ‘hub’ region within the anterior temporal lobes (ATL), allowing the conceptual similarity of items that are semantically similar yet share few surface features, such as ‘flute’ and ‘violin’, to be represented, and making it possible to map between modalities so that we can picture a flute and imagine the sound that it makes from only its name (e.g., Damasio, 1989; Lambon et al., 2010; Patterson et al., 2007; Rogers et al., 2004). This hub and spoke model proposes that both the ATL and modality-specific spokes make a crucial contribution to conceptual representation, and these elements are mutually-constraining through a pattern of interactive-activation.

The spokes are hypothesized to represent the contributions of sensory and motor cortex to conceptual knowledge, as words associated with specific sensorimotor attributes activate corresponding sensorimotor cortex. For example, words denoting actions (e.g., kick) activate the motor system (Postle et al., 2008; Rueschemeyer et al., 2007; Rueschemeyer et al., 2010), while words associated with specific smells (e.g., cinnamon) elicit activation in olfactory cortex (Cerf-ducastel and Murphy, 2004; Gonzalez et al., 2006). Although these neural regions are important for perception and action, they are also recruited during semantic processing to provide meaning to words (Barsalou, 1999; 2008; Martin, 2007; Patterson, et al., 2007; Kiefer and Pulvermuller, 2012).

The proposal that the ATL forms a key semantic “hub” capturing knowledge across different input modalities was initially put forward to account for the pattern of impairment in semantic dementia (SD), in which relatively focal atrophy centered on ATL leads to progressive conceptual degradation across modalities and tasks (e.g., Patterson et al., 2007; Rogers et al., 2015). SD patients are highly consistent in the knowledge they can demonstrate when the same concepts are probed in different ways, suggesting central semantic representations degrade in this condition. Patients with SD have atrophy which increasingly affects inferior frontal and posterior temporal areas, as
Next we used the regions identified in our MVPA analysis as regions of interest in a seed based resting state connectivity analysis to understand the neural networks in which these different regions of the ATL are embedded. We expected the amodal region of ATL to show functional connectivity with regions of cortex that are important in more abstract forms of cognition, e.g., the default mode network, rather than regions important in unimodal sensory processing, such as the auditory and visual cortex. Finally, we used the search tool Neurosynth to decode the most common interpretations of this pattern of functional connectivity in the broader neuroimaging literature.

2. Materials and methods

2.1. Functional experiment

2.1.1. Participants

Twenty participants were recruited from the University of York. One participant’s data was excluded due to excessive motion artifacts, leaving nineteen subjects in the final analysis (10 female; mean age 24.55, range 18–36 years). Participants were native British speakers, right handed and had normal or corrected-to-normal vision. Participants gave written informed consent to take part and were reimbursed for their time. The study was approved by the York Neuroimaging Centre Ethics Committee at the University of York.

2.1.2. Stimuli

Participants were presented with blocks of spoken and written items from three conditions: AUD words denoted auditory features (e.g., loud), VIS words denoted visual features (e.g., shiny) and NON stimuli were meaningless nonwords (e.g., brodic). A block consisted of a sequence of items; participants were asked to pay attention to the meaning of each item, and respond with their left index finger when an out-of-category item was presented (see Fig. 1). For VIS and AUD blocks, half of the out-of-category items were taken from the non-presented feature condition, while the other half were taken from a

![Fig. 1](image-url)
Footnote: Log frequency=log-transformed lemma frequencies from the SUBTLEX database (Brysbaert, New & Keuleers, 2012; http://expsys.agent.ubc.ca/subtlex). Length=number of letters. Age of acquisition (AoA norms; Kuperman et al., 2012). Part of speech also taken from SUBTLEX database. Familiarity, emotional valence and behavioural feature rating (auditory; visual; haptic; taste) were obtained from a behavioural experiment with a separate cohort of participants from the fMRI study. These were scored on a 5-point Likert-scale (1–5).

1 Wilcoxon signed rank tests revealed a significant difference between auditory-feature and visual-feature conditions (p < .001).
2.1.6. Univariate analysis

The condition onset and duration were taken from the first item in each block (after the initial instructions) to the end of the last item. The response to each of the 6 conditions was contrasted against rest. Boxcar regressors for each condition, for each run, in the general linear model were convolved with a double gamma hemodynamic response function (FEAT, FSL). Regressors of no interest were also included to account for head motion within scans. A fixed effect design (FLAME, http://www.fmrib.ox.ac.uk/fsl) was then conducted to average across the four runs, within each individual. Finally, individual participant data were entered into a higher-level group analysis using a mixed effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl) whole-brain analysis.

2.1.7. Multivariate pattern analysis

Parameter estimates were calculated in the same manner as for univariate analyses, for each condition and for each run: in this way, the spatial pattern information entered into the classifier from each condition represented the average response to the 8 exemplars. This method is consistent with previous literature investigating semantic representations (Coutanche & Thompson-Schill, 2012; Fairhall and Kuhl, Rissman, Chun and Wagner, 2011). To ensure that our ROIs had sufficient in each sub-region of the ATL (as signal drop out is most prominent in ventral anterior regions), the tSNR was calculated for the mean signal in each voxel by the standard deviation of the residual function (FEAT, FSL). Regressors of no interest were also included to model were convolved with a double gamma hemodynamic response to each of the 6 conditions was contrasted against rest. Boxcar regressors for each condition, for each run, in the general linear model were convolved with a double gamma hemodynamic response function (FEAT, FSL). Regressors of no interest were also included to account for head motion within scans. A fixed effect design (FLAME, http://www.fmrib.ox.ac.uk/fsl) was then conducted to average across the four runs, within each individual. Finally, individual participant data were entered into a higher-level group analysis using a mixed effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl) whole-brain analysis.

2.2. Resting state fMRI

2.2.1. Participants

This analysis was performed on a separate cohort of 42 healthy participants at York Neuroimaging Centre (13 male; mean age 20.31, range 18–25 years). Subjects completed a 9 minute functional connectivity MRI scan during which they were asked to rest in the scanner with their eyes open. Using these data we examined the resting-state fMRI (rs-fMRI) connectivity of ATL regions that were informative to the semantic feature (aITG) and perceptual classifiers (aSTG) to investigate whether these regions fell within similar or distinct networks. In addition, we investigated the rs-fMRI connectivity of semantic regions within primary sensory cortices that showed significant decoding by the semantic classifiers to examine whether these regions overlap with the connectivity maps of the ATL seeds.

2.2.2. Acquisition

As with the functional experiment, a Magnex head-dedicated gradient insert coil was used in conjunction with a birdcage, radiofrequency coil tuned to 127.4 MHz. For the resting-state data, a gradient-echo EPI sequence was used to collect data from 60 axial slices with an interleaved (bottom-up) acquisition order with the following parameters: TR=3 s, TE=minimum full, volumes=180, flip angle=90°, matrix size=64x64, FOV=192x192 mm, voxel size=3x3x3 mm. A minimum full TE was selected to optimise image quality (as opposed to selecting a value less than minimum full which, for instance, would be beneficial for obtaining more slices per TR). Functional images were co-registered onto a T1-weighted anatomical image from each participant (TR=7.8 s, TE=3 ms, FOV=290 mmx290 mm, matrix size=256 mm x256 mm, voxel size=1 mmx1 mmx1 mm).
2.2.3. Pre-processing

Data were preprocessed using the FSL toolbox (http://www.fmrib.ox.ac.uk/fsl). Prior to conducting the functional connectivity analysis, the following pre-statistics processing was applied to the resting state data; motion correction using MCFLIRT to safeguard against motion-related spurious correlations (Baker et al., 2015; Smallwood et al., 2016; Krieger-Redwood et al. 2016; Davey et al., 2016); slice-timing correction using Fourier-space time-series phase-shifting; non-brain removal using BET; spatial smoothing using a Gaussian kernel of FWHM 6 mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=100 s); Gaussian lowpass temporal filtering, with sigma=2.8 s.

2.2.4. Low level analysis

For our ATL sites we created two spherical seed ROIs, 6 mm in diameter, centered on the co-ordinates of the central voxel in the highest performing spheres in our presentation and semantic searchlight analyses; left aSTG [-54 2 -10] and aITG [-50 -10 -26] respectively (see Table 2). For our sensory semantic regions we created two spherical seed ROIs centered on intracalcarine cortex [-18 -84 4] and planum polare [-48 -12 -4] from the best performing spheres in our semantic searchlight analysis; as these regions showed high performance accuracy on the semantic classifier and fall within primary sensory regions.

The time series of these regions were extracted and used as explanatory variables in a separate subject level functional connectivity analysis for each seed. Subject specific nuisance regressors were determined using a component based noise correction (CompCor) approach (Behzadi et al., 2007). This method applies principal component analysis (PCA) to the fMRI signal from subject specific white matter and CSF ROIs. In total there were 11 nuisance regressors, five regressors from the CompCor and a further 6 nuisance regressors were identified using the motion correction MCFLIRT. These principle components are then removed from the fMRI data through linear regression. The WM and CSF covariates were generated by segmenting each individual’s high-resolution structural image (using FAST in FSL; Zhang et al., 2001). The default tissue probability maps, referred to as Prior Probability Maps (PPM), were registered to each individual’s high-resolution structural image (T1 space) and the overlap between these PPM and the corresponding CSF and WM maps was identified. These maps were then thresholded (40% for the SCF and 66% for the WM), binarized and combined. The six motion parameters were calculated in the motion-correction step during pre-processing. Movement in each of the three Cartesian directions (x, y, z) and rotational movement around three axes (pitch, yaw, roll) were included for each individual.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mask</th>
<th>Cluster Peak</th>
<th>Extended Cluster Regions</th>
<th>Cluster Extent</th>
<th>Z-score</th>
<th>Acc (%)</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semantic Feature</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATL</td>
<td>L Anterior ITG/MTG</td>
<td>L Heschls gyrus, L putamen</td>
<td></td>
<td>478</td>
<td>4.91</td>
<td>61.22</td>
<td>-50</td>
<td>-10</td>
<td>-26</td>
</tr>
<tr>
<td>ATL</td>
<td>R Temporal pole</td>
<td>R Anterior parahippocampal gyrus, R Anterior MTG, R Anterior STG</td>
<td></td>
<td>416</td>
<td>4.58</td>
<td>61.05</td>
<td>42</td>
<td>12</td>
<td>-24</td>
</tr>
<tr>
<td>ATL</td>
<td>L Planum polare</td>
<td>L Heschls gyrus, L Planum temporale</td>
<td></td>
<td>88</td>
<td>3.92</td>
<td>59.53</td>
<td>-48</td>
<td>-12</td>
<td>-4</td>
</tr>
<tr>
<td>ATL</td>
<td>L Intracalcarine cortex</td>
<td>L Lingual gyrus</td>
<td></td>
<td>81</td>
<td>4.26</td>
<td>61.18</td>
<td>-18</td>
<td>-84</td>
<td>4</td>
</tr>
<tr>
<td><strong>Presentation format</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual</td>
<td>L Occipital pole</td>
<td>L Occipital fusiform gyrus, L Inferior lateral occipital cortex</td>
<td></td>
<td>607</td>
<td>4.3</td>
<td>58.57</td>
<td>-16</td>
<td>-92</td>
<td>0</td>
</tr>
<tr>
<td>Visual</td>
<td>L Anterior STG</td>
<td>L Temporal pole, R Anterior STG</td>
<td></td>
<td>66</td>
<td>2.8</td>
<td>58.36</td>
<td>-58</td>
<td>-10</td>
<td>-2</td>
</tr>
</tbody>
</table>

Footnote: Highest decoding accuracy clusters for semantic feature (AUD vs. VIS) and presentation format (spoken vs. written words) analysed separately. Semantic feature classifier was trained on the distinction between spoken AUD vs. spoken VIS and tested on written AUD vs. written VIS (and vice versa). Presentation format classifier was trained on the distinction between written non-words vs. spoken non-words and tested on spoken words vs. written words. Results are thresholded at p < .05 (cluster corrected). L=left, R=right. As well as peak accuracy (reported under the ‘Cluster Peak’ column), the ‘Extended Cluster Regions’ includes all significant regions within each ROI. In addition to the searchlight analyses reported in the table, a further searchlight analysis was run on the distinction between all spoken vs. all written items. This revealed accuracies as high as 99.6% in primary sensory regions and 93.2% in ATL. The unthresholded MVPA maps for each searchlight have been uploaded to the Neurovault database and can be found here http://neurovault.org/collections/1970/.
2.2.5. High level analysis

At the group-level the data were processed using FEAT version 5.98 part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl) and the analyses were carried out using FMRIB's Local Analysis of Mixed Effects (FLAME) stage 1 with automatic outlier detection. The z statistic images were then thresholded using clusters determined by $z > 2.3$ and a (corrected) cluster significance threshold of $p = 0.05$ (Worsley, 2001). No global signal regression was performed.

To investigate the differences between the connectivity maps a fixed effect design (FLAME, http://www.fmrib.ox.ac.uk/fsl) was conducted for each participant to investigate four contrasts; (i) aSTG > aITG seed, (ii) aITG > aSTG seed, (iii) auditory semantic > visual semantic seed and (iv) visual semantic > auditory semantic seed. Individual participant data were then entered into a higher-level group analysis using a mixed effects design (FLAME, http://www.fmrib.ox.ac.uk/fsl) whole-brain analysis. Finally, to determine whether our ATL seeds connectivity maps overlap with the connectivity maps of the sensory semantic seeds we calculated the number of overlapping voxels for our two ATL sites and the sensory semantic connectivity maps.

2.3. Resting state decoder

To allow quantitative inferences to be drawn on the functional neural activity identified through our seed based correlational analyses we performed an automated meta-analysis using NeuroSynth (http://neurosynth.org/decode; Yarkoni et al., 2011). This software computed the spatial correlation between each ATL component mask and every other meta-analytic map ($n=11406$) for each term/concept stored in the database (e.g., semantic, language, memory, sensory). The 15 meta-analytic maps exhibiting the highest positive correlation and negative correlation for each sub-system mask were extracted, and the term corresponding to each of these meta-analyses is shown in Fig. 4. The font size reflects the size of the correlation (ranging from $r=0.10$ to 0.45 for positive correlations and $r=-0.05$ to $-0.2$ for negative correlations, in increments of 0.05). This allows us to quantify the most likely reverse inferences that would be drawn from these functional maps by the larger neuroimaging community.

3. Results

3.1. Behavioural results

Accuracy and reaction times (RT) were calculated for each participant ($n=19$) for the catch trials in each experimental condition. Results showed that all participants paid attention to the words as indicated by a mean accuracy above 80% for all experimental conditions (spoken AUD = 80.63% ± 15.33, spoken VIS = 88.12% ± 4.86, spoken NON=85.62% ± 11.47, written AUD=83.12% ± 19.01, written VIS=86.25% ± 13.52, written NON=88.75% ± 5.45). A chi-square test of independence revealed that accuracy did not significantly differ across the six experimental conditions ($\chi^2(5)=6.09, p=.303$) or across spoken and written input ($\chi^2(1)=301.3$, RTs differed significantly between modality-input ($t(59)=7.36, p<.001$), but not semantic-category within each modality (spoken: F(2,38)=9.2, written: F(2,38)=0.74, ns). In line with previous findings (Booth et al., 2002; Cohen et al., 2004), participants were significantly faster at responding to written than spoken stimuli. Furthermore, there was no difference in RT between AUD, VIS and NON items within each presentation modality, suggesting that the experimental conditions were well matched at the behavioural level within our stimuli subset.

3.2. Searchlight analysis

3.2.1. Semantic feature classifier

The format-independent searchlight classifier, trained on the distinction between visual and auditory features in one presentation modality and tested on this distinction in the other modality, was run in three separate masks (ATL; primary auditory cortex and primary visual cortex). All results reported are above chance levels (50%, cluster corrected $p<.05$). The searchlight analysis within the ATL mask revealed a left hemisphere cluster that could decode semantic information across modalities in aMTG and aITG (see Fig. 3, Table 2). Additionally, right hemisphere clusters were revealed in anterior parahippocampal gyrus and temporal pole (TP). The searchlight analysis within the primary auditory mask revealed a cluster in planum.
Finally, the primary visual cortex mask revealed a cluster in intracalcarine cortex that could decode semantic content (see Fig. 5, Table 2). The classifier results for the primary visual cortex mask revealed an extensive cluster of voxels in occipital pole (see Fig. 5; Table 2).

3.2.2. Perceptual classifier

The classifier that was trained on the distinction between spoken and written non-words and tested on the distinction between these presentation modalities for words, was also run in three separate masks (ATL; primary auditory cortex and primary visual cortex). All results reported are above chance levels (50%, cluster corrected p < .05). Within the ATL, anterior portions of STG, extending into temporal pole, were able to decode between presentation formats (see Fig. 3; Table 2). The classifier results for the primary auditory cortex mask revealed an extensive cluster of voxels that could classify perceptual information in Heschl’s Gyrus, planum temporale and superior temporal gyrus (see Fig. 4; Table 2). The classifier results for the primary visual cortex mask revealed an extensive cluster of voxels in occipital pole (see Fig. 5; Table 2).

To explicitly determine whether the aITG and aSTG were differentially able to classify the modality of presentation and the meaning of the stimulus, we conducted a 2×2 repeated-measures ANOVA in which we compared the prediction accuracies for each classifier output for each significant cluster. This revealed three significant effects. First, a main effect for classifier type (presentation format vs. semantic classifier; F(1,18)=36.76, p < .001). Second, a significant main effect of region (aSTG vs. aITG; F(1,18)=79.71, p < .001). Critically, we also found a significant interaction between classifier type and ATL region (F(1,18)=1087.51, p < .001). Post-hoc tests revealed a significant difference between aSTG and aITG for the presentation format classifier, with aSTG performing significantly better than aITG (t(18)
presented location and size of each of these
Figs. 7 (occipital pole and intracalcarine cortex). All four sites showed strong
and semantic feature type within primary auditory cortex (planum
We also examined regions that could decode modality of presentation
(aSTG) was sensitive to modality (spoken=written words). In contrast,
aITG showed substantial overlap with limbic and DMN
The current study used multiple imaging methods to identify
regions in the anterior temporal lobe (ATL) and primary sensory
regions that showed the pattern expected for the semantic hub of the
and provides evidence that these regions are situated within distinct
and provide evidence that these regions are situated within distinct
temporal lobe. The unthresholded maps from the whole-brain search-
clusters across primary auditory cortex, primary visual cortex and
condition from the pairs of clusters that were able to decode semantic
monkeys, pMTG and LIFG. These differences resemble resting
state differences for aSTG and vATL reported by Jackson et al. (2016),
helping to validate the functional dissociation we observed using
MVPA.

4. Discussion

The searchlight results revealed that in ATL, primary auditory
cortex and visual cortex, distinct regions were able to decode semantic
feature type and presentation modality. As an additional complemen-
tary analysis, the percentage signal change was extracted for each
condition from the pairs of clusters that were able to decode semantic
feature type and modality of presentation in ATL, visual cortex and
auditory cortex (generating six analyses; see Fig. 6). A 6mm sphere
was centered at the peak MVPA accuracy in each of these sites (see Table 2).
The ventral ATL region (encompassing aITG and aMTG, decoding
feature type) showed deactivation across all four conditions, and the
degree of deactivation was sensitive to meaning (auditory>visual features)
but not input modality (spoken>written words). In contrast, aSTG (which decoded presentation modality) was sensitive to modality
(spoken>written) but not meaning (auditory=visual features). Thus,
univariate analyses also revealed a functional dissociation within ATL.
We also examined regions that could decode modality of presentation
and semantic feature type within primary auditory cortex (planum temporale and planum polare respectively) and primary visual cortex
(occipital pole and intracalcarine cortex). All four sites showed strong
effects of input modality in univariate analyses across both feature
types. In addition, the intracalcarine cortex showed greater activity to
words that denoted a visual property (e.g., bright) whereas planum
polare showed greater activation to words that denoted an auditory
property (e.g., loud). This effect of meaning in primary visual and
auditory areas was only seen when the words were presented in the
complementary input modality: primary visual cortex responded more
to visual features when written words were presented, while primary
auditory cortex responded more to auditory features when spoken
words were presented. Thus, aITG was unique in showing a pattern
across both multivariate and univariate analyses consistent with the
predictions for an amodal ‘hub’: i.e., sensitivity to meaning and
insensitivity to presentation modality.

3.3. Univariate analysis

The current study used multiple imaging methods to identify
regions in the anterior temporal lobe (ATL) and primary sensory
regions that showed the pattern expected for the semantic hub of the
hub and spokes model (Patterson et al., 2007). In an fMRI study,
participants listened to or viewed words that referred to either visual or
auditory features (e.g., **bright** or **loud**). Multivoxel pattern analysis
(MVPA) revealed a dissociation between (i) anterior inferior temporal
gyrus (aITG), which could classify semantic categories relating to
feature type (e.g., auditory features like “loud” as being different from
visual features like “bright”) across auditory and visual inputs and (ii)
*superior temporal gyrus (aSTG), which was sensitive to input
modality across meaningful and meaningless items. This dissociation
within ATL was further supported by univariate contrasts and patterns of
resting state connectivity: aSTG showed a stronger response to
spoken than written inputs and was functionally coupled to an
auditory-motor network (somatosensory network; Yeo et al., 2014),
while aITG was insensitive to input modality and showed substantial
connectivity with regions in the default mode network and limbic
network, plus some overlap with visual regions (see Jackson et al.
(2016), for similar findings).

Our findings make an important contribution to our understanding
of the neural basis of semantic cognition in three ways: (1) We provide evidence that conceptual knowledge, extracted from different modal-
ities of input across many learning experiences, is represented within
ventral portions of ATL which act as a ‘hub’ (Patterson et al., 2007;
Rogers et al., 2004). (2) Across converging methods, we observe a
functional dissociation between ventral and superior portions of ATL
and provide evidence that these regions are situated within distinct
large-scale cortical networks. (3) Responses in primary visual and
auditory cortex confirm the contribution of these ‘spoke’ regions to
semantic processing.

According to the hub and spoke model (Patterson et al., 2007),
conceptual knowledge depends on the co-activation of spoke regions that convey information about specific unimodal and multimodal features of concepts, and an ATL hub which integrates these features to form amodal conceptual representations that are independent of specific sensory input. Studies of patients with semantic dementia (SD) provided the original motivation for this proposal yet neuropsychological methods are not especially well-suited to the precise localization of amodal conceptual representations given the widespread atrophy in this condition. Nevertheless, the degree of semantic impairment correlates with hypometabolism in ventral rather than superior portions of ATL across patients (Mion et al., 2010), suggesting that ventral ATL could be the critical substrate for amodal knowledge. Relevant evidence is also provided by univariate fMRI analyses of the ATL response to verbal comprehension tasks in healthy participants, which show multiple peak responses in both ventral ATL and aSTG, often to the same contrasts (Binney et al., 2010; Hoffman et al., 2015; Visser and Ralph, 2011). Semantic matching and naming tasks have also shown multiple peak responses in the ATL with the more superior ATL region being involved in object naming and the more ventral region in semantic matching (Sanjuán et al., 2015). Furthermore, the differential patterns of functional connectivity across ATL regions have been observed by both Jackson et al., (2016) and Pascual et al., (2015).

Our findings therefore add to existing knowledge by showing a dissociable response in these two regions: only the ventral ATL site showed a pattern consistent with the representation of conceptual information, since it was able to classify responses according to semantic category (i.e., feature type, not input modality). In univariate analyses, this aITG site also showed deactivation (arguably due to the use of rest rather than an active baseline; Visser et al., 2010; Humphreys et al., 2015) for both auditory and visual feature types, irrespective of whether these words were spoken or written—and the magnitude of this deactivation was greater for visual than auditory features. Finally, this site showed stronger functional connectivity at rest with the default mode and limbic systems, as expected for a region implicated in amodal conceptual processing. Therefore, our combination of functional and resting state methods provides novel converging evidence that anterior ventral temporal areas allow different sensory representations to be integrated to form ‘amodal’ conceptual representations (particularly for auditory features, see limitations below).

Previous studies have used MVPA to explore the neural basis of semantic processing, and have identified a conceptual response in ATL using classification of stimuli within a single presentation modality (Coutanche and Thompson-Schill, 2014; Peelen and Caramazza, 2012). Other studies, examining semantic cognition across modalities of

<table>
<thead>
<tr>
<th>6mm ROIs centered on peak MVPA clusters</th>
<th>% signal change in semantic classifier ROIs</th>
<th>% signal change in modality classifier ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior Temporal Lobe</td>
<td>AUD</td>
<td>VIS</td>
</tr>
<tr>
<td>aITG</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>aSTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary auditory cortex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary visual cortex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planum polare</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planum temporale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Cortex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracalcarine cortex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occipital pole</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 6. The first column shows 6 mm ROIs centered on the peak MVPA results from the searchlight analyses (shown in Figs. 3–5) for semantic classifier in magenta and modality classifier in cyan, for each of our three masks (anterior temporal lobe, primary auditory cortex and primary visual cortex) projected in blue. The centre for these ROIs are as follows; aITG seed [-50 -10 -26], aSTG seed [-58 -10 -2], planum polare [-48 -12 -4], planum temporale [-58 -24 8], intracalcarine cortex [-16 -84 4] and occipital pole [-16 -92 0]. The second column shows the univariate percent signal change for each of our four conditions within the semantic (magenta) ROI. The third column shows the univariate percent signal change for each of our four conditions with the modality (cyan) ROI. Grey bars show the results for auditory-feature words (e.g., ‘loud’) and white bars show the results for visual-feature words (e.g., ‘bright’). * indicates a significant difference between auditory-features and visual-features within a modality (i.e., spoken auditory-features and spoken visual-features; p < .05). ** indicates a significant difference between spoken and written presentation format (p < .001). The unthresholded univariate maps for each condition have been uploaded to the Neurovault database and can be found here http://neurovault.org/collections/1970/. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
presentation (Devereux et al., 2013; Fairhall and Caramazza, 2013; Man et al., 2015), have largely not observed effects in ATL. An exception is a recent crossmodal MVPA study, investigating Dutch-English bilinguals (Correia et al., 2014). The research tested whether patterns of activity related to the distinction between spoken nouns in one language (e.g., "horse" vs. "duck" in English) could accurately predict the same distinction in the other language (e.g., "paard" vs. "eend" in Dutch). Consistent with our findings, the cross-language classifier revealed a significant cluster in the left ATL. This largely fell within mid-superior temporal pole rather than the more ventral region we identified in our analysis, perhaps because aSTG is an important interface between semantic processing and other aspects of language.

Analyses of resting state connectivity from the ATL regions that were able to classify input modality (aSTG) and semantic feature type (aITG) revealed that these two sites lie within distinct large-scale functional networks. A similar dissociation between the resting state connectivity of ventral ATL and anterior STG was recently reported by Jackson et al., (2016), providing further evidence for the validity of the functional dissociation in ATL that we observed using MVPA. To quantify the interpretation of the functional connectivity of the aSTG and aITG connectivity maps, we performed a decoding analysis using automated MRI meta-analytic software NeuroSynth (see Fig. 8). Meta-analytic decoding of these spatial maps revealed that our aSTG connectivity map correlated with terms related to language (e.g., sentence, comprehension) and auditory processing (e.g., speech, sound) whilst anti-correlating with other modality information (e.g., visual, spatial) and memory (e.g., working memory, episodic). In contrast, the aITG connectivity map correlated with terms related to memory (e.g., semantic, autobiographical) and social processes (e.g., theory of mind, social cognition) terms, whilst anti-correlating with modality-specific (e.g., ventral visual, motor, spatial) and executive terms (e.g., maintenance, demands). This is consistent with previous findings that relate aSTG to speech comprehension, language and sensory processing (Patterson and Ralph, 1999; Jobard et al., 2007; Scott and Johnsrude, 2003; Scott et al., 2000; Spitsyna et al., 2006) and aITG to semantic processing but not sensory experience (Patterson et al., 2007; Visser et al., 2010). Furthermore, the differences in function across temporal areas as revealed by the Neurosynth database seem to align with differences in the white-matter terminations (see Bajada et al., 2016). These findings confirmed associations between (i) the network anchored in the aSTG and auditory processing and speech perception, plus (ii) the aITG network and more abstract domains (such as social cognition, theory of mind, or mental states).

Thus, the putative semantic ‘hub’ in ventral ATL was functionally coupled to aspects of cortex that specialize in forms of stimulus-independent higher order cognition, including angular gyrus (AG) and posterior and anterior areas on the medial surface that correspond to...
the midline core of the so-called default mode network (DMN) (see also Hurley et al., 2015). This network is known to be deactivated by input (Raichle et al., 2001) and is thought to be crucial in the generation of cognitive states that do not rely on sensory information for their mental content (for a review see Andrews-Hanna et al., 2014). Tasks which are associated with the default mode network include those that depend on episodic memory, semantic processing, mental state attribution as well as states of spontaneous thought studied under the rubric of mindwandering / daydreaming (Spreng et al., 2009; Raichle, 2015). Although previous literature has shown that connectivity to the AG may not be due to shared semantic processing (Humphreys et al., 2015). Therefore, as many cognitive states that involve the DMN are stimulus-independent in nature, their association with ventral ATL both in terms of functional connectivity and their meta-analytic decoding is consistent with the view that this region supports semantic processing across different input modalities and may form conceptual representations that are not tied to a specific input modality (see Margulies et al., 2016). In contrast, aSTG showed greater functional connectivity with auditory and motor regions and this spatial map was associated with auditory processing and language tasks, as opposed to amodal tasks, in the meta-analytic decoding. Therefore, our combination of functional and resting state methods provides novel converging evidence that anterior ventral temporal areas allow di-different sensory representations to be integrated to form ‘amodal’ conceptual representations.

As discussed, the hub and spoke model (Ralph et al., 2010; Patterson et al., 2007; Rogers et al., 2004) makes novel predictions about the contribution of the ATL to amodal conceptual knowledge, but it also anticipates an important role for modality-specific ‘spoke’ regions in visual and auditory cortex, in line with many influential accounts of semantic processing (Damasio, 1989; Martin, 2007; Hurley et al., 2015; Meteyard et al., 2012; Pulvermüller, 2013). Furthermore, the involvement of both hub and spoke regions in semantic representations has been shown using TMS (Pobric et al., 2010). In line with this view, MVPA revealed regions that responded to meaning in both ventral parts of ATL (putative ‘hub’) and in primary visual and auditory regions (putative ‘spokes’). In addition, even though the putative ‘spoke’ regions (i.e., voxels sensitive to meaning) were adjacent to areas that coded for input modality, the specific voxels that could classify meaning and input modality were largely different. These findings do not readily support traditional ‘strong’ embodied accounts that equate semantic representations with traces of perceptual/motor experience (for a review, see Meteyard et al. (2012)) since this would suggest a greater degree of overlap between the results of these two classifiers. While our data suggests that sensory systems appear to play a critical role in the representation of meaning, they also suggest that perceptual experience and imagery generated as part of semantic retrieval may be distinguishable on the basis of differences in the patterns of activity in sensory cortex.

One potential limitation of our study is that we did not observe evidence that aITG responds to both auditory and visual semantic features in the univariate contrasts: this site showed deactivation for both feature types that was greater for visual features. Thus, the strongest evidence for the aITG as an amodal hub is provided by the MVPA results and our meta-analytic decoding of this region’s pattern of distinct functional connectivity, and not the univariate analyses. Our design was optimized for decoding rather than univariate effects – as we focused on obtaining the maximum number of blocks for MVPA and did not employ a high-level non-semantic baseline which would have

Table 3 Coordinates of peak clusters in the resting-state connectivity analyses.

<table>
<thead>
<tr>
<th>Seed Region</th>
<th>Cluster</th>
<th>Cluster Extent</th>
<th>Z-score</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>aSTG</td>
<td>Increased Correlation</td>
<td>L. aSTG</td>
<td>15745</td>
<td>12.3</td>
<td>-54</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Temporal pole</td>
<td>12970</td>
<td>9.24</td>
<td>52</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cingulate Gyrus</td>
<td>7618</td>
<td>7.02</td>
<td>-12</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>Reduced Correlation</td>
<td>L. Cuneal cortex</td>
<td>26667</td>
<td>6.19</td>
<td>-20</td>
<td>-74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Superior frontal gyrus</td>
<td>4128</td>
<td>4.69</td>
<td>20</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Middle frontal gyrus</td>
<td>2259</td>
<td>4.53</td>
<td>32</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Lateral occipital cortex, inferior</td>
<td>1457</td>
<td>5.46</td>
<td>-46</td>
<td>-70</td>
</tr>
<tr>
<td>aITG</td>
<td>Increased Correlation</td>
<td>L. aITG/MTG</td>
<td>20324</td>
<td>13.1</td>
<td>-50</td>
<td>-10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Frontal pole</td>
<td>2899</td>
<td>7.22</td>
<td>-10</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Occipital fusiform gyrus</td>
<td>1981</td>
<td>4.49</td>
<td>26</td>
<td>-82</td>
</tr>
<tr>
<td></td>
<td>Reduced Correlation</td>
<td>Postcentral gyrus</td>
<td>3725</td>
<td>4.44</td>
<td>0</td>
<td>-54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Frontal pole</td>
<td>2717</td>
<td>5.07</td>
<td>42</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. IFG, pars triangularis</td>
<td>2118</td>
<td>5.17</td>
<td>-46</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Cingulate gyrus</td>
<td>1276</td>
<td>4.44</td>
<td>12</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Angular gyrus</td>
<td>783</td>
<td>4.39</td>
<td>-40</td>
<td>-50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Superior parietal lobule</td>
<td>769</td>
<td>3.94</td>
<td>-30</td>
<td>-48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L. Middle frontal gyrus</td>
<td>724</td>
<td>4.72</td>
<td>-28</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R. Middle frontal gyrus</td>
<td>626</td>
<td>4.16</td>
<td>30</td>
<td>12</td>
</tr>
</tbody>
</table>

Footnote: The table shows peak clusters in the resting-state connectivity analysis from two seed regions; aSTG and aITG. Results are thresholded at p < .01 (cluster corrected). L=left, R=right.
allowed us to recover semantic activation in ATL for both auditory and visual features from a contrast (Humphreys et al., 2015). Since we found that aITG responds more to auditory features (words such as “loud”) than visual features (words such as “bright”), it remains unclear whether aITG reflects the meanings of auditory features alone, or both feature types equally. Future studies might allow these possibilities to be disentangled using a high-level baseline with which both feature types can be compared (e.g. Jackson et al., 2015).

5. Conclusion

Collectively, our findings from both pattern classification and resting-state connectivity provide converging evidence that sub-regions of the ATL support different aspects of semantic processing. Anterior ITG and MTG capture meaning independent of input modality, consistent with the fact that semantic dementia patients (who have multimodal semantic impairment) have considerable atrophy in this same region of ATL (Binney et al., 2010; Galton et al., 2001). In contrast, aSTG exhibited a degree of modality specificity: this structure, which is known to be important for understanding speech and environmental sounds, does not fulfill the criteria for an amodal semantic hub. Finally, the current results provide evidence for modality-specific spokes regions within the vicinity of primary auditory and visual cortex (intralacaral cortex and planum polare respectively). However, the specific voxels that could classify between each condition (presentation format and semantic feature) were largely different. These findings challenge traditional embodied accounts (Pulvermüller, 2005) that attempt to equate semantic representations with traces of perceptual/motor experience, and instead support the view that the richness of semantic cognition arises at least in part from abstraction away from specific input modalities in ventral regions of the anterior temporal lobe.

Funding

The research was supported by BBSRC grant BB/J006969/1. Jefferies was supported by a grant from the European Research Council (SEMBIND – 283530). The publication was part-funded by a grant from the John Templeton Foundation, “Prospective Psychology Stage 2: A Research Competition” to Martin Seligman. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation. The authors declare no competing financial interests.

Acknowledgements

We would like to thank Tim Andrews for his advice about the experimental design.

References


for 30,000 English words. Behav. Res. Meth. 44 (4), 978–990.
Lambon Ralph, M.A., Sage, K., Jones, R., Maybery, E., 2010. Coherent concepts are
2717–2722.
Levenshtein, V.I., 1965. Binary codes with correction for deletions and insertions of the
MA, 911–920.
representations for sight, sound, and touch. Human. brain Mapp. 36 (9),
3629–3640.
Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G.,
Jefferies, E., et al., 2016. Situating the default-mode network along a principal
gradient of macroscale cortical organization. Proc. Natl. Acad. Sci. 113 (44),
12574–12579.
58, 25–45.
Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y.,
T., et al., 2010. What the left and right anterior fusiform gyri tell us about semantic
memory. Brain 133, 3256–3268.
Pascual, B., Masdeu, J.C., Hollenbeck, M., Makris, N., Insausti, R., Ding, S.L., Dickerson,
Patterson, K., Lambon Ralph, M.A., 1999. Is a picture worth a thousand words? Evidence
from concept definitions by patients with semantic dementia. Brain Lang. 70 (3),
309–335.
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. Nat. Rev. Neurosci. 8,
964–975.
20 (10), 964–968.
word meaning representations in cytoarchitectonically defined primary and
Neurosci. 6, 578–582.
Rachle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L.,
Rice, G.E., Lambon Ralph, M.A. Hoffman, P. The roles of left versus right anterior
temporal lobes in conceptual knowledge: An ALE meta-analysis of 97 functional
Rogers, T.T., Lambon Ralph, M.A., Garrett, P., Rosset, S., McColland, J.L., Hodges,
J.R., Patterson, K., 2004. Structure and deterioration of semantic memory: a
neuropsychological and computational investigation. Psychol. Rev. 111 (1),
205–235.
Rogers, T.T., Patterson, K., Jefferies, E., Lambon Ralph, M.A., 2015. Disorders of
representation and control in semantic cognition: effects of familiarity, typicality and
novel correlates of processing verbs with motor stems. J. Cogn. Neurosci. 19,
855–865.
Sajjan, A., Hope, T.M., Jones, O.F., Prejawa, S., Oberhuber, M., Guerin, J., Price, C.J.,
2015. Dissociating the semantic function of two neighbouring subregions in the
Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of
Smallwood, J., Karapanagiotidis, T., Ruby, F., Medea, B., de Caso, I., Konishi, M.,
Jefferies, E., 2016. Representing representation: integration between the temporal
lobe and the posterior cingulate influences the content and form of spontaneous
thought. Plos one 11 (4), e0152272.
143–155.
language stream systems in the human temporal lobe. J. Neurosci. 26 (28),
7528–7536.
memory, prospection, navigation, theory of mind, and the default mode: a
anterior temporal lobes and semantic memory clarified: novel evidence from a
temporal gyrus and the ventral anterior temporal area are crucial for multimodal
semantic processing: distortion-corrected ePI mri evidence for a double gradient of
temporal lobe and left anterior superior temporal gyrus to semantic processes. J.
Cogn. Neurosci. 23 (10), 3121–3141.
introduction to methods 14, 251–270.
scale automated synthesis of human functional neuroimaging data. Nat. Methods 8,
665–670.
and overlap of functional connectivity networks in the human cerebral cortex.
NeuroImage 0, 212–227.
Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M.,
Fisch, B., et al., 2011. The organization of the human cerebral cortex estimated by
hidden Markov random field model and the expectation-maximization algorithm.