Barking up the right tree: Univariate and multivariate fMRI analyses of homonym comprehension

Paul HOFFMAN*

Andres TAMM

School of Philosophy, Psychology & Language Sciences, University of Edinburgh, UK

* Correspondence to:
Dr. Paul Hoffman
School of Philosophy, Psychology & Language Sciences,
University of Edinburgh,
7 George Square, Edinburgh, EH8 9JZ, UK
Tel: +44 (0) 131 650 4654
Email: p.hoffman@ed.ac.uk

Acknowledgements

The study was funded by grants from the Carnegie Trust for the Universities of Scotland (70743) and Moray Endowment Fund. PH was also supported by the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged.

Abstract

Homonyms are a critical test case for investigating how the brain resolves ambiguity in language and, more generally, how context influences semantic processing. Previous neuroimaging studies have associated processing of homonyms with greater engagement of regions involved in executive control of semantic processing. However, the precise role of these areas and the involvement of semantic representational regions in homonym comprehension remain elusive. We addressed this by combining univariate and multivariate fMRI analyses of homonym processing. We tested whether multi-voxel activation patterns could discriminate between presentations of the same homonym in different contexts (e.g., *bark* following *tree* vs. *bark* following *dog*). The angular gyrus and ventral anterior temporal lobe, regions implicated in semantic representation but not previously in homonym comprehension, showed this meaning-specific coding, despite not showing increased mean activation for homonyms. Within inferior frontal gyrus (IFG), a key site for semantic control, there was a dissociation between BA47, which also showed meaning-specific coding, and BA45, which discriminated more generally between semantically related and unrelated word pairs. Unlike the representational regions, IFG effects were goal-dependent, only occurring when the task required semantic processing, in line with a top-down control function. Finally, posterior middle temporal cortex showed a hybrid pattern of responses, supporting the idea that it acts as an interface between semantic representations and the control system. The study provides new evidence for context-dependent coding in the semantic system and clarifies the role of control regions in processing ambiguity. It also highlights the importance of combining univariate and multivariate neuroimaging data to fully elucidate the role of a brain region in semantic cognition.

Keywords: ambiguity; context; comprehension; semantic control; hub-and-spoke.

Introduction

Language is littered with ambiguity. Most words have multiple semantic interpretations whose relevance depends on context (e.g., leather *belt* vs. asteroid *belt*). Often the various uses for a word appear to share a common semantic core; this is known as polysemy. This is not the case for homonyms like *bark,* however. *Bark* can refer either to the sound of a dog or to the covering of a tree, but these meanings have no semantic properties in common; they just happen to share the same phonological and orthographic form. Only around 7% of English words are homonyms (Rodd, Gaskell, & Marslen-Wilson, 2002) but because their ambiguity is so distinct and well-defined, they represent a valuable test case for investigating how semantic processing is influenced by context more generally.

Psycholinguistic studies indicate that when we process homonyms in natural language, both meanings are briefly activated and compete for selection (Duffy, Morris, & Rayner, 1988; Vitello & Rodd, 2015). This competition is typically resolved within a few hundred milliseconds and the most contextually-appropriate meaning selected to guide ongoing comprehension (Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982). Neuroimaging studies have implicated the left inferior frontal gyrus (IFG) in this meaning selection process. In listening tasks, left IFG shows stronger activation for sentences when they contain homonyms (Mason & Just, 2007; Rodd, Davis, & Johnsrude, 2005; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007) and homonyms elicit greater IFG activation than unambiguous words in lexical decision and semantic judgement tasks (Bedny, McGill, & Thompson-Schill, 2008; Grindrod, Garnett, Malyutina, & den Ouden, 2014; Whitney, Grossman, & Kircher, 2009).

These findings are well-explained by the controlled semantic cognition framework, which accounts for semantic processing in terms of interactions between semantic representation and control systems, supported by distinct neural networks (Hoffman, McClelland, & Lambon Ralph, 2018; Jefferies & Lambon Ralph, 2006; Lambon Ralph, Jefferies, Patterson, & Rogers, 2017). On this view, IFG is involved in top-down control over the activation and selection of semantic knowledge represented elsewhere in the cortex. Homonyms are assumed to greater demands on this system because they require retrieval and selection of the contextually appropriate meaning (Noonan, Jefferies, Corbett, & Lambon Ralph, 2010). Different functions have been ascribed to different subregions within IFG. The

more anterior portion (BA47) is thought to support controlled retrieval of semantic knowledge, when the required knowledge is weakly associated with the stimulus and hence not activated automatically by spread of activation (Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; Badre & Wagner, 2007). In contrast, posterior IFG (BA45/44) is implicated in resolution of competition between active lexical-semantic representations (Nagel, Schumacher, Goebel, & D'Esposito, 2008; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). At present it is unclear whether the engagement of IFG during homonym processing reflects one or both of these processes, because both subregions show similar responses to ambiguous words (for discussion, see Vitello & Rodd, 2015).

Left posterior temporal regions also show increased fMRI activation in response to semantically ambiguous words, frequently centred on the posterior middle temporal gyrus (pMTG) (Rodd et al., 2005; Rodd, Vitello, Woollams, & Adank, 2015; Zempleni et al., 2007). The interpretation of these effects is more contested. One view is that pMTG is involved in control processes similar to those ascribed to IFG (Jefferies, 2013; Noonan, Jefferies, Visser, & Lambon Ralph, 2013; Whitney, Jefferies, & Kircher, 2011). Other researchers have proposed that pMTG is involved in representation of lexical-semantic information (or access to these representations), which would also be taxed during homonym comprehension (Bedny et al., 2008; Lau, Phillips, & Poeppel, 2008; Tyler, Cheung, Devereux, & Clarke, 2013). One fMRI study used multiple priming of homonyms (e.g., game-dance-ball) to attempt to differentiate between these accounts (Whitney et al., 2011). Both IFG and pMTG were sensitive to variation in the control demands of the task but not to the number of meanings that were retrieved on each trial, which appears inconsistent with a representational account. However, since ambiguity-related pMTG effects are weaker and less spatially consistent than those in IFG (Vitello & Rodd, 2015), there is less certainty over the role of this region.

The studies described thus far have used univariate activation contrasts to implicate IFG and pMTG in homonym comprehension. It is generally assumed that contrasts of this type index differences in the degree to which stimuli engage the processes or representations supported by the region (Taylor, Rastle, & Davis, 2013). Therefore, IFG and pMTG regions may show greater activation for homonyms because their multiple meanings necessitate greater engagement of the semantic control processes supported by these regions. What about regions implicated in representation of semantic knowledge? Current theories implicate anterior temporal and inferior parietal regions in semantic representation (Binder & Desai,

2011; Lambon Ralph et al., 2017). These areas do *not* typically show increased activation for homonyms, suggesting that the presence of ambiguity does not place greater metabolic demands on brain regions that encode semantic knowledge. Does this mean that these regions are not involved in meaning disambiguation? This seems unlikely. Multivariate fMRI studies have shown that activation patterns in anterior temporal and inferior parietal cortex vary according to the semantic properties of words and objects, supporting the idea that these regions code information about semantic content (Bruffaerts et al., 2013; Devereux, Clarke, Marouchos, & Tyler, 2013; Fairhall & Caramazza, 2013; Peelen & Caramazza, 2012). It is therefore likely that elements of this network change their multivariate response to homonyms depending on the currently-relevant meaning, even in the absence of mean activation changes. Despite the recent burgeoning of multivariate fMRI studies, this hypothesis remains untested.

In the present study, we used multi-voxel pattern analysis (MVPA) to investigate how neural activation patterns vary in homonym comprehension, focusing on three representational brain areas in addition to the control-related regions discussed earlier. Two of these regions, the left angular gyrus (AG) and left lateral anterior temporal lobe (ATL), have been particularly implicated in semantic representation at a multi-word level. These regions show increased neural responses to coherent combinations of words, for example coherent adjective-noun phrases (e.g., *loud car*) or meaningful sentences (Bemis & Pylkkänen, 2012; Graves, Binder, Desai, Conant, & Seidenberg, 2010; Humphries, Binder, Medler, & Liebenthal, 2006; Price, Bonner, Peelle, & Grossman, 2015). Both regions have therefore been associated with combinatorial semantic processing, i.e., the extraction of a global meaning from a series of words (Bemis & Pylkkänen, 2012; Price et al., 2015; Vandenberghe, Nobre, & Price, 2002). The combinatorial semantic processing associated with both AG and lateral ATL would seem to be critical to homonym comprehension, where it is necessary to integrate the homonym with prior context in order to determine the appropriate semantic interpretation.

In contrast, the ventral portion of the ATL (inferior temporal and fusiform gyri) is strongly implicated in multimodal semantic processing at the single word/concept level (Lambon Ralph et al., 2017). Ventral ATL shows robust activation to individual words as well as multi-word combinations (Humphreys, Hoffman, Visser, Binney, & Lambon Ralph, 2015) and a series of multivariate neuroimaging studies have shown that activation patterns in this region discriminate object properties and word meanings (Clarke & Tyler, 2014; Coutanche,

2013; Peelen & Caramazza, 2012). Theoretical accounts hold that the ventral ATL acts as a semantic hub that binds and integrates different linguistic and perceptual elements of experience to form coherent concepts (Lambon Ralph et al., 2017; Patterson, Nestor, & Rogers, 2007; Rogers et al., 2004). Some accounts of ATL function posit that its representations must be context-independent, in order for conceptual knowledge to generalise appropriately across contexts (Binney, Parker, & Lambon Ralph, 2012; Lambon Ralph et al., 2017). Other models suggest that it is advantageous for the hub to be sensitive to context, to make use of semantic information present in the distributional statistics of language (Hoffman et al., 2018). Empirically, however, the degree to which lexical-semantic representations in the ventral ATL are independent of context remains an unanswered question. Homonyms provide a useful test case here because the same lexical item takes on radically different meanings in different contexts.

Only one previous fMRI study has investigated neural representation of homonyms in the ventral ATL. Musz and Thompson-Schill (2017) presented participants with homonyms in sentences that primed their dominant or their subordinate meanings. They then compared the neural patterns elicited by the same homonym in the two different contexts. In the ventral ATL, they found that the similarity between dominant and subordinate patterns was predicted by the polarity of the homonym (i.e., the degree to which the dominant meaning occurs more often than subordinate one in natural language). When the dominant meaning was much more common than the subordinate one, the patterns in ventral ATL were more similar to one another. One interpretation of this result is that ventral ATL representations *do* vary as function of homonym meaning, but that highly dominant meanings are activated to some degree even when they are irrelevant. This may have caused the dominant and subordinate patterns to resemble one another for highly polarised homonyms.

In the present study, we investigated neural responses to balanced homonyms in which neither meaning was highly dominant over the other. We used univariate and multivariate fMRI to investigate patterns of neural engagement and information coding across the semantic network during homonym processing. By combining these distinct sources of information about homonym processing, we aimed to assess (1) the degree to which IFG sub-regions (BA45, BA47) and pMTG support the control and selection of meanings and (2) the degree to which activation patterns in AG, lateral and ventral ATL vary according to homonym meaning. Participants were presented with sequential word pairs in which one

meaning of a homonym was primed (e.g., *tree-bark* vs. *dog-bark*). This allowed us to assess whether activation patterns in each brain region could successfully discriminate between the alternative meanings of each homonym. On other trials, the prime was related to neither meaning (e.g., *letter-bark*), allowing us to test whether activity patterns discriminated the presence of absence of a semantic relationship. We also manipulated the task participants performed. In semantic runs, they decided whether the two words were semantically related; in phonological runs, they made syllable judgements in which the meaning of the words was irrelevant. This allowed us to assess the degree to which processes engaged during homonym comprehension occur automatically, in the absence of an explicit comprehension goal.

Method

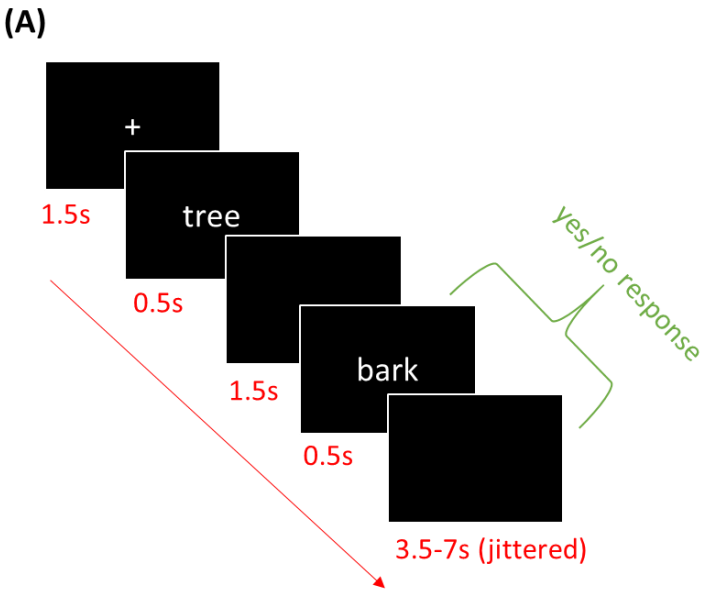
Participants: 24 native English speakers took part in the study (17 female; mean age = 22.3; sd = 4.2). All were classified as right-handed using the Edinburgh Handedness Inventory (Oldfield, 1971) and none reported dyslexia or any history of neurological illness. All provided informed consent. Data from one participant was excluded because they performed at chance when making semantic judgements about the meanings of homonyms (all other participants scored >85%). Neuroimaging and behavioural data for all participants are available on the Open Science Foundation repository: https://osf.io/ut3f9/.

Stimuli: We investigated semantic processing of ten target words. Five of these were homonyms with two distinct meanings (*bark, calf, cell, pupil, seal*). The other five were unambiguous words matched to the homonyms for word length, frequency and concreteness (*coal, menu, monk, poet, wizard*). Each target word was paired with 12 different primes (see Figure 1B and 1C for examples). Eight of these were selected to have a strong semantic relationship with the target while the other four were unrelated in meaning. For the homonyms, half of the related primes were related to each of the target's meanings, allowing us to investigate activity associated with opposing meanings of the same word. Semantic relatedness ratings were collected for all prime-target pairs from a group of undergraduate students who did not participate in the main experiment. Homonyms and unambiguous targets were matched for mean prime-target relatedness, as well as the frequency,

concreteness and length of their primes (see Supplementary Table 1). For the homonyms, care was also taken to ensure that the primes for each meaning had equally strong associations with the target. Stimuli were divided into four sets, for presentation in different scanning runs. In each set, each target appeared three times, once with an unrelated prime and twice with related primes. For homonyms, the two related primes in each set primed opposing meanings (see Figure 1B).

Procedure: Participants completed eight runs of scanning. In each run, they were presented with one set of 30 prime-target pairs. In half of the runs, they made semantic judgements about the prime-target pairs (are these words related in meaning?). In the other half, they made phonological judgements (do these words contain the same number of syllables?). The order of tasks was counterbalanced over participants. Manual responses were made using the left and right hands, with the mapping of these to response options counterbalanced over participants. Prior to entering the scanner, participants were warned that some of the targets had more than one meaning and that their primes could relate to either meaning. They were given brief definitions of the two meanings for each homonym (e.g., *Bark can mean the noise made by a dog or the covering of a tree*). They practiced both tasks before entering the scanner.

The timeline for a single trial is shown in Figure 1A. Each trial began with a fixation cross presented for 1.5s. This was followed by the prime, which appeared for 0.5s. After a 1.5s delay, the target was presented for 0.5s. Participants were instructed to respond as quickly as possible upon seeing the target. Trials were separated by a mean inter-trial interval of 5s (jittered between 3.5s and 7s). Participants saw each set of trials once in the semantic task and once in the phonological task, with the order of presentation of the sets counterbalanced. Within runs, trials were presented in a different random order for each participant.



(B)

	Unrelated	Related 1	Related 2
Run 1	bun, bark	oak, bark	labrador, bark
Run 2	rag, bark	wood, bark	woof, bark
Run 3	letter, bark	forest, bark	dog, bark
Run 4	town, bark	tree, bark	growl, bark

(C)

	Unrelated	Related 1	Related 2
Run 1	pillow, coal	mining, coal	smoke, coal
Run 2	hockey, coal	oil, coal	furnace, coal
Run 3	speed, coal	chimney, coal	flame, coal
Run 4	curtain, coal	stove, coal	burning, coal

Figure 1: Experimental design (A) Timeline for a single trial. (B) Prime-target pairs for a homonym target. (C) Prime-target pairs for an unambiguous target.

Image acquisition and processing: Images were acquired on a 3T Siemens Prisma scanner using a 32-channel head coil. A dual-echo protocol was employed in which gradient-echo EPI images were simultaneously acquired at two TEs (13ms and 35ms) and a mean of the two echo series was computed during preprocessing (Halai, Parkes, & Welbourne, 2015). This approach improves signal quality in the ventral ATLs, which typically suffer from susceptibility artefacts (Ojemann et al., 1997). The TR was 1.8s and images consisted of 60 slices with a 100 x 100 matrix and voxel size of 2.4mm isotropic. Multiband acceleration with a factor of 2 was used and the flip angle was 74°. Eight runs of 158 volumes (284s) were acquired. A high-resolution T1-weighted structural image was also acquired for each participant using an MP-RAGE sequence with 0.8mm isotropic voxels, TR = 2.62s, TE = 4.5ms.

Images were preprocessed and analysed using SPM12. Preprocessing steps consisted of slice-timing correction, spatial realignment and unwarping using a fieldmap, and normalisation to MNI space using Dartel (Ashburner, 2007). For univariate analyses, images were smoothed with a kernel of 8mm FWHM. Data were treated with a high-pass filter with a cut-off of 128s and the eight runs were analysed using a single general linear model. For each run, a single regressor modelled presentation of prime words, each with a duration of

0s. Targets were modelled with four regressors per run corresponding to the four experimental conditions (related-ambiguous, related-unambiguous, unrelated-ambiguous, unrelated-unambiguous). Each target was modelled with a duration of 2s. Covariates consisted of the six motion parameters and their first-order derivatives, as well as mean signal in white matter and CSF voxels.

Our main analyses focused on anatomical regions of interest (ROI) defined below. Contrast estimates were extracted from these regions using marsbar (Brett, Anton, Valabregue, & Poline, 2002) and analysed with ANOVA. At a whole-brain level, we also tested for the main effect of task, and, in each task, effects of relatedness and ambiguity. These analyses were thresholded at a voxel level of $p<0.005$ and corrected for multiple comparisons at the cluster level using SPM's random field theory ($p<0.05$ corrected for familywise error).

Regions of interest: Our main analyses focused on left-hemisphere anatomical regions of interest (ROIs), selected a priori based on their involvement in semantic representation or control. These are shown in Figure 4A. Five of the six ROIs were defined using probability distribution maps from the Harvard-Oxford brain atlas (Makris et al., 2006), including all voxels with a >30% probability of falling within the following regions:

BA47: the pars orbitalis region, with voxels more medial than x=-30 removed to exclude medial orbitofrontal cortex (Hoffman, 2019)

BA45: the pars triangularis region

pMTG: the temporo-occipital part of the middle temporal gyrus

Lateral ATL: the anterior division of the superior and middle temporal gyri

Ventral ATL: the anterior division of the inferior temporal and fusiform gyri

The final ROI covered the angular gyrus and included voxels with a >30% probability of falling within this region in the LPBA40 atlas (Shattuck et al., 2008). A different atlas was used in this case because the AG region defined in the Harvard-Oxford atlas is particularly small and does not include parts of the inferior parietal cortex typically implicated in semantic processing.

Multivariate pattern analysis: For MVPA, normalised functional images were smoothed with a 4mm FWHM kernel (Hendriks, Daniels, Pegado, & Op de Beeck, 2017). Each run was analysed with a separate GLM in which each of the 30 targets was modelled with a separate regressor (with a single regressor again modelling the presentation of primes). T-maps were

generated for each target presentation and t-values from the voxels in each anatomical ROI were extracted for use in decoding analyses. The Decoding Toolbox (v3.997) was used for these analyses (Hebart, Görgen, & Haynes, 2015).

Decoding analyses were performed separately on the four runs of the meaning task and the four runs of the phonological task. Classifiers were trained to discriminate between two classes of stimuli, using a support vector machine (from the LIBLINEAR library). To ensure independence of training and test data, we used a cross-validated leave-one-run out approach, in which the classifier was trained on data from three scanning runs and tested on the remaining run. The regularisation parameter *C*, which determines the classifier's tolerance to misclassifications, was allowed to vary between 10^{-4} and 10^{3}. The optimum *C* value for each training run was selected using leave-one-run-out nested cross-validation (Hebart et al., 2015).

Two forms of classification analysis were performed.

Classifier 1: Related vs. unrelated trials. All targets were included in this analysis, coded according to their semantic relatedness, and the classifier was trained to discriminate related from unrelated trials (see Figure 2A). This analysis tested whether the neural patterns in each region reliably coded the presence or absence of a semantic relationship, irrespective of which particular word was being processed. The univariate analysis revealed that some regions showed differences in overall activation between related and unrelated trials. So that the classifier could not use these mean activation differences to discriminate the two classes, we mean-centred the activation patterns for each trial prior to classification (Coutanche, 2013).

Classifier 2: Meaning1 vs. meaning2 for homonyms. Only the related trials with homonym targets were used in this analysis. We took all related trials with the same homonym target and trained a classifier to discriminate which meaning was primed (see Figure 2B). We repeated this process for each of the five homonyms in turn. This analysis therefore tested for meaning encoding at a word-specific level. It determined whether the neural patterns in each region reliably coded which meaning of the homonym was accessed.

To determine whether classification in each region was significantly better than expected by chance, second-level random-effects analyses were performed. Given the binomial nature of the data (each classification attempt at test could be correct or incorrect), a binomial generalised linear mixed model was estimated on the classification outcomes for

each region (0=incorrect classification; 1=correct classification), including an intercept and random effects of participant. Such a model has an intercept of zero when correct and incorrect outcomes are equally likely, after taking individual participant variation into account (Jaeger, 2008). Therefore, the classifier was deemed to have achieved above-chance performance if the intercept was significantly greater than zero (tested using a likelihood ratio test against a model with no intercept).

(A) CLASSIFIER 1

	Class 1: Unrelated	Class 2: Related
Train	rag, bark sap, wizard volume, seal …	oak, bark magic, wizard leak, seal …
Test	letter, bark nerve, wizard path, seal …	tree, bark witch, wizard arctic, seal …

(B) CLASSIFIER 2

	Class 1: Meaning 1	Class 2: Meaning 2
Train	oak, bark wood, bark forest, bark	labrador, bark woof, bark dog, bark
Test	tree, bark	growl, bark

Figure 2: Illustration of multivariate classifier analyses Figure shows example stimuli used to train and test each classifier in a single cross-validation fold. (A) Classifier 1 was trained to distinguish related from unrelated trials. All targets were included in train and test sets. Figure shows a subset of all trials that were included in the analysis. (B) Classifier 2 was trained to distinguish between the two meanings of each homonym. This analysis was performed separately for each homonym and the results averaged.

Results

Behavioural performance: Mean accuracy and RT in each condition are reported in Table 1. Generalised linear mixed effects models were used to assess effects of task, ambiguity and meaning relatedness, and their interactions, on performance. Models included random intercepts and slopes for participants and run and trial order were included as covariates. For the accuracy data, there was a main effect of task ($\chi^2 = 18.5$, $p < 0.001$), as overall performance was better for the meaning task, and relatedness ($\chi^2 = 19.4$, $p < 0.001$), since more correct responses were given when prime and target were semantically unrelated. There were also

significant interactions of task with both relatedness (χ^2 = 26.0, p < 0.001) and ambiguity (χ^2 = 6.3, p = 0.012). Separate analyses performed on each task indicated that relatedness influenced performance on the meaning task (χ^2 = 15.6, p < 0.001) but not on the syllable task (χ^2 = 0.4, p = 0.51). This was expected, since the relatedness of the word pairs was irrelevant for the syllable task. In contrast, ambiguity only influenced accuracy on the syllable task (χ^2 = 4.3, p = 0.038).

Analyses were performed on RT data following log-transformation to reduce skew. The main model indicated main effects of task (t(21.2) = 4.2, p < 0.001) and relatedness (t(24.5) = 7.1, p < 0.001). On average, participants were faster to respond in the meaning task and when the prime and target were semantically related. The task manipulation interacted with relatedness (t(23.3) = 6.3, p < 0.001) and ambiguity (t(22.1) = 5.3, p < 0.001). Separate analyses for each task revealed that, for the meaning task, participants were faster to respond to related word pairs (t(22.3) = 7.4, p < 0.001), while no such effect was present in the syllable task (t(2332) = 0.02, p = 0.98). In the meaning task, participants were slower to respond to homonym targets (t(22.5) = 4.5, p < 0.001). Conversely, they were faster to respond to homonyms in the syllable task (t(25.8) = 4.0, p < 0.001). This result is consistent with previous studies showing that lexical ambiguity has a negative effect when people make semantic judgements but is beneficial in other lexical processing tasks (Hino, Lupker, & Pexman, 2002; Hoffman & Woollams, 2015).

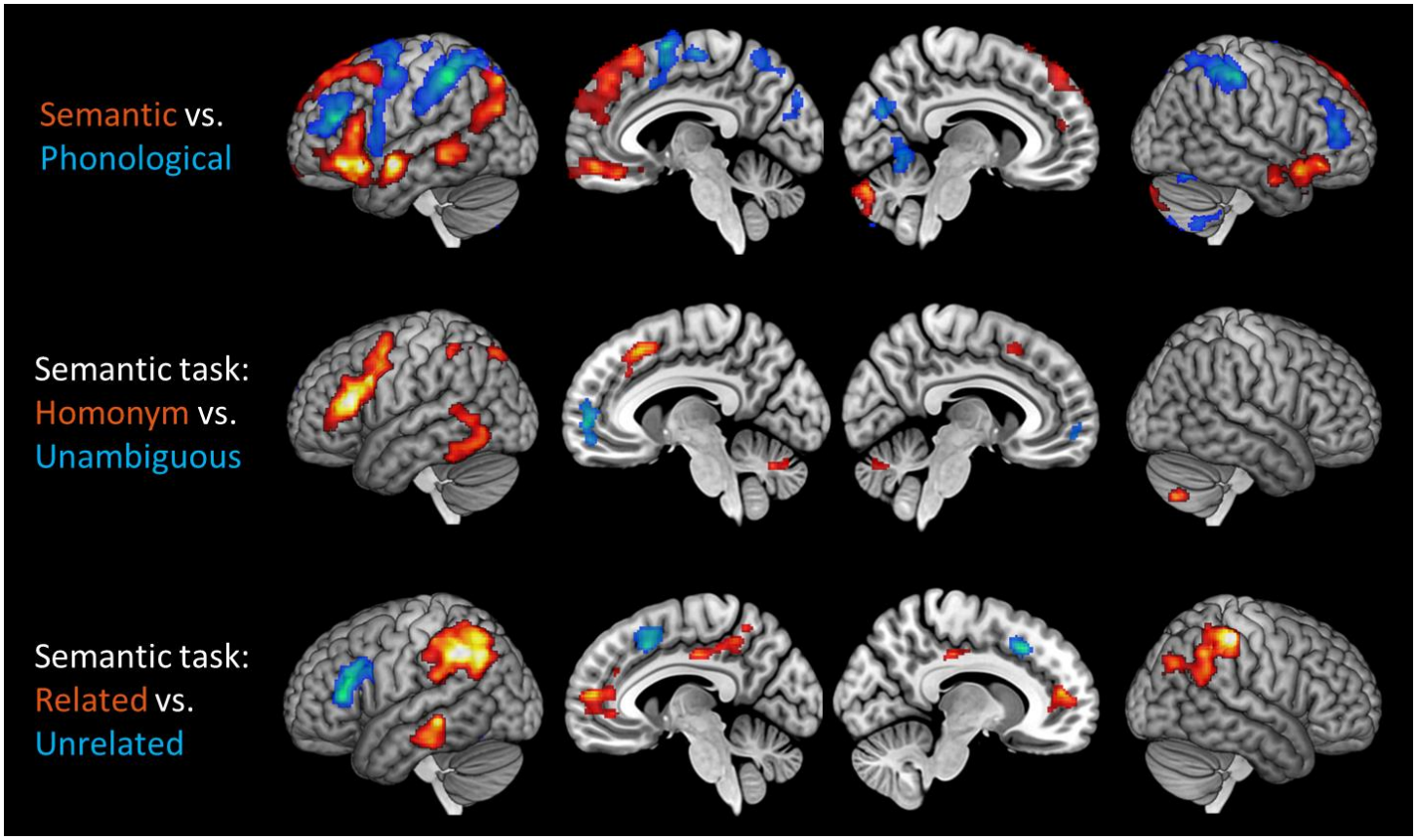
Table 1: Behavioural performance

	Accuracy		RT	
	Semantic task	Phonological task	Semantic task	Phonological task
Homonyms, related trials	88.3 (7.7)	89.9 (13.9)	1238 (294)	1462 (385)
Homonyms, unrelated trials	97.6 (4.7)	86.7 (15.8)	1511 (395)	1439 (362)
Unambiguous, related trials	92.9 (9.6)	85.6 (14.7)	1164 (278)	1532 (392)
Unambiguous, unrelated trials	98.0 (3.3)	85.9 (16.6)	1412 (361)	1548 (437)

Standard deviations are reported in parentheses.

Univariate analyses: Whole-brain contrasts for the semantic vs. phonological task are shown in Figure 3. The semantic task produced greater activation in a range of predominately left-

lateralised cortical regions implicated in semantic processing, including IFG, anterior and posterior temporal regions and AG. Greater activation for the phonological task was observed in frontal and parietal regions associated with phonological processing, working memory and cognitive control. Effects of relatedness and ambiguity in the semantic task are also presented in Figure 3. As expected, greater activation for homonyms was predominately observed in left prefrontal cortex and posterior temporal cortex. Semantically unrelated trials also engaged left prefrontal regions to a greater extent, while more activation on related trials was found in a number of default mode network regions, including bilateral AG, posterior cingulate and ventromedial prefrontal cortex. In the phonological task, no effects of ambiguity were found but greater engagement for semantically related word pairs was found in bilateral AG, bilateral middle frontal gyrus and right IFG (see Supplementary Figure 1).



<u>Figure 3: Whole-brain univariate activation contrasts</u> Images are shown at a threshold of p<0.05, corrected for multiple comparisons at the cluster level.

Effects of ambiguity and relatedness within our ROIs are shown in Figure 4B (Supplementary Figure 2 shows activations in each condition relative to rest). We first performed a 6 x 2 x 2 x 2 ANOVA (region x task x ambiguity x relatedness) on the ROI data to determine whether the effects of our experimental manipulations varied across regions. They did: region interacted with task, ambiguity and relatedness (all $F > 7.4$, $p < 0.001$) and there were additional 3-way interactions between region, task and ambiguity and region, task and relatedness ($F > 6.2$, $p < 0.001$). We therefore performed separate 2 x 2 ANOVAs (ambiguity x relatedness) on the data from each ROI, split by task. As shown in Figure 4B, greater activation on homonym trials was observed in BA45 only (though in pMTG, the expected effect of ambiguity during the semantic task was just short of statistical significance at $p = 0.058$). In contrast, lateral ATL showed less activation for homonyms than for unambiguous words during the semantic task. Divergent effects of semantic relatedness were also present. Both IFG regions showed greater activation for the more difficult unrelated trials, but only when relatedness was task-relevant (i.e., during the semantic task). AG, conversely, was more engaged during related trials and this effect was not task-dependent. Finally, we conducted a direct comparison of the effects for each task in BA45 vs. BA47, using 2 x 2 x 2 ANOVAs. During the semantic task, BA45 showed larger effects of the ambiguity manipulation (ambiguity x region interaction: $F(1,22) = 22.1$, $p < 0.001$) and the relatedness manipulation (relatedness x region interaction: $F(1,22) = 5.72$, $p = 0.03$). There were no such effects for the phonological task, although BA45 did show much greater overall activation than BA47 for this task ($F(1,22) = 57.3$, $p < 0.001$).

<u>Multivariate pattern analysis:</u> Decoding accuracies for the first MVPA classifier (related vs. unrelated trials) are shown in Figure 4C. Neural patterns in BA45 and pMTG reliably signalled the presence or absence of a semantic relationship, but only during the semantic task when this distinction was task-relevant. In contrast, patterns in AG discriminated between related and unrelated trials during both tasks. No coding of trial status was found in BA47 or the lateral or ventral ATL.

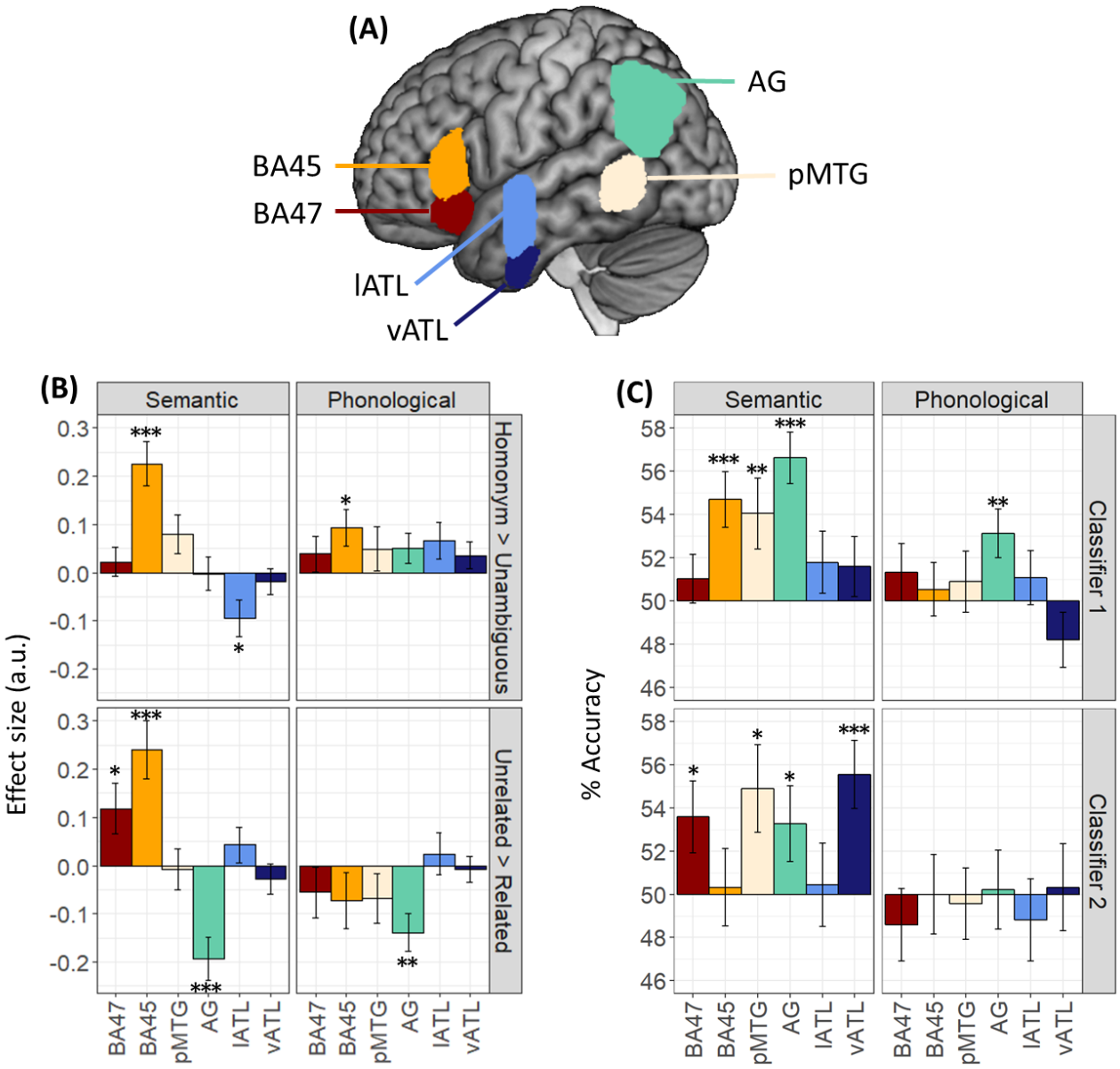


*Figure 4: Region of interest results (A) Locations of anatomically-defined ROIs. (B) Univariate contrasts of ambiguity and relatedness for each task. (C) Classification accuracies for decoding analyses applied to each task. Bars indicate SEM; * = p<0.05; ** = p<0.01; *** = p<0.001.*

Performance of the second classifier (meaning 1 vs. meaning 2 for homonyms) is also shown in Figure 4C. No regions showed above-chance decoding during phonological processing; however, during the semantic task BA47, pMTG, AG and ventral ATL all exhibited above-chance classification. This indicates that the neural patterns in these regions reliably

signalled which meaning of the homonym was relevant to the trial, suggesting that these areas code the opposing meanings of homonyms differently.

These results suggest that, during the semantic task, different regions in the semantic network coded different types of information about the stimuli. To determine whether this was the case, we entered classification accuracies into a mixed effects model with fixed effects of classifier analysis (1 vs. 2) and region, plus random effects (intercepts and slopes) of participants. This analysis revealed an interaction between classifier and region ($\chi^2 = 15.6$, $p < 0.001$), confirming that differences between the two classifiers varied as a function of region. We repeated this analysis using only data from BA45 and B47 and the region x analysis interaction was again significant ($\chi^2 = 6.48$, $p = 0.01$). This confirms that the two IFG subregions code different information about each trial, with BA45 coding semantic status (related or unrelated) at a general level, while BA47 codes word-specific information about the relevant meaning.

Discussion

We used univariate and multivariate fMRI to investigate how different elements of the semantic neural network process the variable meanings of homonyms. Our principal finding was that various areas in left frontal, temporal and parietal cortices exhibited activation patterns that reliably predicted which of the homonym's meanings was relevant to the trial. Although previous studies have found that neural coding patterns vary according to the word being comprehended, here we found that systematic variation in the patterns elicited by the same word in different contexts. A distinct set of regions coded at a more general level for the semantic status of each trial. Importantly, multivariate stimulus coding was observed in the absence of univariate activation differences between trial types and vice versa. Our results are supportive of a broad distinction between regions that represent semantic knowledge and those that regulate and control its use. They also provide new insights into distinct roles played by different elements of the semantic network when processing ambiguity. In what follows, we first discuss interpretation of effects in individual regions before considering implications for more general network-level accounts of the semantic system.

Inferior frontal gyrus

A major aim of the study was to clarify the role of left IFG in resolving semantic ambiguity. Although the entire IFG region is implicated in semantic control, researchers have proposed a distinction between anterior IFG (BA47), which is thought to play a key role in controlling the activation and retrieval of knowledge from semantic memory, and the more posterior portion (BA45), which is involved in selecting task-relevant representations (Badre et al., 2005; Badre & Wagner, 2007; Nagel et al., 2008; Thompson-Schill et al., 1997). Despite this hypothesised functional dissociation, previous studies have provided little evidence for distinct roles of BA47 vs. BA45 in processing semantic ambiguity (Vitello & Rodd, 2015). Here, however, we found that these areas showed divergent profiles in both their levels of overall neural engagement and in the type of stimulus information coded in their neural patterns. In univariate analyses, only BA45 showed greater activation for homonyms compared with unambiguous targets, while both regions were more activated for trials containing no semantic relationship. Importantly, MVPA analyses also produced divergent results. In BA45, neural patterns reliably coded whether a semantic relationship was present. In contrast, BA47 neural patterns coded item-specific semantic information, shifting their patterns of activity according to which of meaning was task-relevant.

These divergent effects suggest different roles in regulating the semantic activation produced by homonyms. Results in BA47 support the assertion that this area provides top-down support to semantic retrieval, which is critical when automatic stimulus-driven activity does not identify the required semantic representation. This account explains greater activation on semantically unrelated trials, because these require a sustained retrieval effort in order to thoroughly search semantic memory and discount the possibility that a semantic relationship is present. It also explains why we observed this effect in the semantic but not the phonological task, since a controlled search for meaning is not engaged when participants are not motivated to process the stimulus at a semantic level. Our MVPA results also suggest that BA47 was engaged in meaning-specific retrieval processes, since the neural patterns in this region varied depending on the particular meaning that was relevant to the trial. In contrast, patterns did not vary according to the required behavioural response (related vs. unrelated), suggesting that BA47 is involved in identifying trial-relevant semantic information but not in using this information to determine a response.

We did not observe greater univariate activation in BA47 when participants processed homonyms. This is surprising as these words might have been expected to place greater demands on controlled retrieval in order to activate the correct aspect of meaning. There are two potential explanations for this. First, the appropriate meaning was primed prior to presentation of the homonym itself, which may have reduced the need for control. Second, we used balanced homonyms in which both meanings were similarly frequent in language and therefore both relatively accessible in the semantic system. Other studies that have used biased homonyms have found greater BA47 activation when subordinate (infrequent) meanings are retrieved (Mason & Just, 2007; Whitney et al., 2009), in line with a greater need for controlled retrieval (Hoffman et al., 2018; Noonan et al., 2010).

In BA45, greater activation for unrelated trials and for homonyms is consistent with the more general selection and competition resolution mechanisms attributed to this region. On semantically unrelated trials, participants must reject and inhibit any irrelevant semantic information accessed as participants search for a meaningful connection between the words. Greater BA45 engagement for homonyms is also readily explained in terms of competition between their two alternative meanings. In addition, related trials for homonyms induce additional competition between potential response options, since any activation of the currently-irrelevant meaning would direct participants towards making an "unrelated" response (Pexman, Hino, & Lupker, 2004).

MVPA results indicate that activation patterns in BA45 were attuned to the correct behavioural response for the trial, but not to which word-specific elements of meaning were retrieved. This result suggests that BA45 is less intimately involved in processing the semantic properties of the stimuli per se, and more in using the semantic information to determine an appropriate behavioural response. Thus, overall our data suggest that BA47 and BA45 play different roles in processing ambiguous words, in line with the established distinction between retrieval and selection functions (Badre & Wagner, 2007). Our data are also compatible with the more general assertion that anterior IFG is specifically engaged by semantic processing while posterior IFG plays a more general role in various aspects of controlled language processing (Gough, Nobre, & Devlin, 2005; Krieger-Redwood, Teige, Davey, Hymers, & Jefferies, 2015); and with functional and structural connectivity data indicating that BA47 shows strong connectivity with anterior temporal regions linked specifically with semantic processing while more posterior regions (BA44/45) are also

connected with frontoparietal networks involved in domain-general cognitive control (Jackson, Hoffman, Pobric, & Lambon Ralph, 2016; Jung, Cloutman, Binney, & Lambon Ralph, 2016; Xiang, Fonteijn, Norris, & Hagoort, 2009).

Anterior temporal lobe

The hub-and-spoke model holds that the ATL acts as a hub for integrating multi-modal information into conceptual representations, with the ventral portion of the ATL forming the heteromodal centre of this representational region (Lambon Ralph et al., 2017; Rice, Hoffman, & Lambon Ralph, 2015). The role of this region in comprehending homonyms has rarely been investigated; this is a significant lacuna because there are differing theoretical perspectives on the degree to which these conceptual representations should be sensitive to context (Binney et al., 2012; Hoffman et al., 2018; Lambon Ralph et al., 2017). Our data provide empirical support for the context-sensitivity of semantic representations in the ventral ATL, since it exhibited distinct neural patterns for the same word depending on which of its meanings was currently relevant. This supports the view that the hub uses recent linguistic context to shape its representations (Hoffman et al., 2018), allowing it to leverage the considerable semantic information present in the statistical structure of natural language (Andrews, Vigliocco, & Vinson, 2009; Landauer & Dumais, 1997; Mikolov, Chen, Corrado, & Dean, 2013).

How does this result fit with other theories that emphasise the context-independence of the hub? Such theories have typically defined context in terms of a top-down representation of the current task or goal, arguing that the ATL must be insensitive to these demands in order to acquire unbiased knowledge about the statistical structure of the environment (Binney et al., 2012; Lambon Ralph et al., 2017). This is a somewhat different idea to the bottom-up priming of meaning that we focus on in the present study. Indeed, our data suggest that ventral ATL *is* relatively insensitive to task requirements and current goals. Patterns in this region did not code for the semantic relatedness of trials, even when this was critical to the task being performed. Nor did it show any univariate effects of ambiguity or relatedness, indicating that the demand characteristics of different trial types did not influence its engagement. Thus, the present data suggest that processing in ventral ATL is relatively "insulated" from top-down goal-related context, but at the same time is influenced by the bottom-up context provided by recent experience.

In contrast to ventral ATL, we found few effects in the more lateral portion of the ATL. This region showed no coding in MVPA analyses, which is surprising given its established role in verbal semantic cognition and in combinatorial semantic processing in particular (Baron & Osherson, 2011; Bemis & Pylkkänen, 2012; Humphries et al., 2006; Westerlund & Pylkkänen, 2014). Of course, there are a number of possible explanations for these null results. It may be that our region of interest did not cover the lateral regions most engaged by the task, or that it encompassed two different functional regions with distinct patterns of coding. At a univariate level, lateral ATL showed greater engagement for unambiguous words. This result is consistent with previous findings of greater engagement in this region for concepts that can be combined more easily (Hoffman, Binney, & Lambon Ralph, 2015; Teige et al., 2019).

Angular gyrus

Angular gyrus presented a different pattern of effects to those in ATL and IFG. Like BA47 and ventral ATL, its activation patterns discriminated between the meanings of individual homonyms. However, unlike these regions, its patterns also coded at a more general level for semantically related vs. unrelated trials. Furthermore, it displayed increased engagement for related trials (the opposite effect to IFG regions) and, unlike IFG regions, these effects were present in both the phonological and semantic tasks. Before setting out our preferred interpretation of these findings, it is worth ruling out an alternative account. As a core element of the default mode network, AG frequently displays task-related deactivation in many cognitive domains, which increases with task difficulty (Humphreys et al., 2015; Mckiernan, Kaufman, Kucera-Thompson, & Binder, 2003). This non-specific disengagement has been proposed to account for effects of semantic manipulations in some previous studies, rather than genuine involvement in semantic processing (Hoffman et al., 2015; Humphreys et al., 2015; Lambon Ralph et al., 2017). This was not the case here, however. In the semantic task, related trials were easier than unrelated trials (in terms of reaction time). However, this was not the case in the phonological task, yet relatedness effects were also observed in AG during this task. Moreover, the fact that multivariate activity patterns were reliably sensitive to item-level stimulus properties suggests that the AG is actively engaged in processing the stimuli.

As most effects in AG were observed independently of task, it seems that this region is not involved in controlled semantic processing. Instead, it appears to be engaged by the

processing of coherent conceptual combinations in a more automatic fashion (Davey et al., 2015; Humphreys & Lambon Ralph, 2014). Our results suggest that AG represents semantic content at multiple levels, coding for both the presence of a semantic relationship and for the particular meaning that is currently relevant. The precise role of AG in semantic processing is disputed. Some theories posit that AG acts as a semantic hub coding event-related or thematic semantic knowledge (Binder & Desai, 2011; Mirman, Landrigan, & Britt, 2017; Schwartz et al., 2011). Others have suggested that AG serves as a short-term buffer for recent multimodal experience (Humphreys & Lambon Ralph, 2014). On the latter view, its function is not specific to semantic cognition but is required in some semantic tasks, particularly when context must be used to constrain semantic processing. Both of these accounts are consistent with our data. Related word pairs are more likely to activate representations of a coherent event than unrelated pairs, which could explain greater AG engagement on these trials. The different meanings of the homonyms would typically elicit different event representations, resulting in different activation patterns. However, similar results would be expected if this region was involved in short-term maintenance of the target word in its particular context.

Posterior middle temporal gyrus

Our final aim was to clarify the role of pMTG in processing semantic ambiguity. pMTG frequently shows increased engagement when homonyms are processed (Rodd et al., 2005; Rodd et al., 2015; Zempleni et al., 2007). In the present study, the effect of homonymy was not significant in our pMTG ROI (p = 0.058), though whole-brain analysis did reveal greater activation for homonyms in this general anatomical region. Some studies have implicated pMTG in semantic control functions (Jefferies, 2013; Noonan et al., 2013; Whitney et al., 2011) while others have associated it with lexical-semantic representation (Bedny et al., 2008; Lau et al., 2008; Tyler et al., 2013). Our results are not wholly consistent with either view. Neural patterns in pMTG coded the semantic status of the trial, but only when this was task-relevant. In this sense, its response was similar to that of BA45, and suggests a general role in using semantic information to determine a behavioural response. However, unlike the IFG regions, engagement of pMTG was not increased for the more demanding semantically unrelated trials. In addition, pMTG activation patterns discriminated between the two meanings of homonyms, which indicates coding at the item-specific level (akin to BA47). This complex set of findings is perhaps best explained in terms of a hybrid role for pMTG, whereby

it is somewhat influenced by task demands but also by the nature of the semantic content being accessed. This fits well with other accounts claiming that pMTG is a functional nexus linking executive/semantic control networks with semantic representational regions including ATL and AG (Davey et al., 2015).

Conclusions and future directions

The present study has revealed a complex set of responses to homonyms within the left-hemisphere semantic network. Ventral ATL and AG both showed systematic variation in activation patterns as a function of homonym meaning, suggesting that the semantic information coded in these regions is sensitive to context. Univariate effects in these regions were independent of task, in line with representational system that process stimulus information independently of task demands. In contrast, IFG regions showed variation in engagement as a function of task, indicating a goal-directed role in manipulation and evaluation of semantic information. However, we found that different subregions of IFG coded different forms of information about the stimuli, suggesting that the more anterior BA47 is engaged in top-down control over the retrieval of specific semantic information while BA45 is resolving semantic competition in order to determine a behavioural response.

Importantly, we found that considering univariate and multivariate effects in combination provided important additional information about the functions of specific regions. For example, the MVPA analyses revealed very similar effects in ventral ATL and BA47: both regions decoded the currently-relevant meaning but not the semantic status of the trial, implicating them in coding and activating the specific semantic properties of the stimuli being presented. However, these regions dissociated in the univariate analyses, with only BA47 showing greater engagement on unrelated trials. Changes in engagement are thought to reflect the degree to which a stimulus draws on a cognitive process supported by the region (Coutanche, 2013). In this case, it appears that the process supported by BA47 is more demanding when no semantic relationship is present, and we have argued that this process is likely to be controlled semantic retrieval. In contrast, the similar levels of activation for related/unrelated and for homonyms/unambiguous trials in ventral ATL suggests that processes supported by this region are engaged equally under all conditions. This is consistent with a more passive representational role which is engaged equally by comprehension under all circumstances.

Finally, while multiple regions showed meaning-specific coding for homonyms in the semantic task, no regions showed this effect when participants made phonological decisions about the words. Does this result indicate that the brain does not disambiguate homonyms under these conditions? Not necessarily. All of our ROIs showed lower levels of engagement for the phonological task relative to the semantic task. Thus, it seems that when participants are focused on performing phonological judgements, in which semantic information is unhelpful, neural resources are directed away from the semantic system. One corollary of this is that signal in these regions is much weaker, and may have been insufficient for the classifier to predict meaning at above-chance level. In other words, it is possible that there were still subtle activation shifts as a function of meaning during phonological processing, but not at a level that we were able to reliably detect with fMRI. Another possibility is that disambiguating neural information is briefly activated upon stimulus processing but is not sustained. The temporal resolution of fMRI is such that brief changes in activity are unlikely to be detected. Studies using multivariate decoding methods applied to EEG or MEG data may be valuable in providing more information about the time-course of semantic activation during homonym comprehension.

References

Andrews, M., Vigliocco, G., & Vinson, D. (2009). Integrating Experiential and Distributional Data to Learn Semantic Representations. *Psychological Review, 116*(3), 463-498.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage, 38*(1), 95-113. doi:DOI 10.1016/j.neuroimage.2007.07.007

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*(6), 907-918.

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901.

Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *Neuroimage, 55*(4), 1847-1852. doi:10.1016/j.neuroimage.2011.01.066

Bedny, M., McGill, M., & Thompson-Schill, S. L. (2008). Semantic adaptation and competition during word comprehension. *Cerebral Cortex, 18*, 2574-2585.

Bemis, D. K., & Pylkkänen, L. (2012). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, bhs170.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*(11), 527-536. doi:DOI 10.1016/j.tics.2011.10.001

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.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. J. J. o. N. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *33*(47), 18597-18607.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*(3), 904-911.

Clarke, A., & Tyler, L. K. J. J. o. N. (2014). Object-specific semantic coding in human perirhinal cortex. *34*(14), 4766-4775.

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us? *Cognitive, Affective, & Behavioral Neuroscience, 13*(3), 667-673.

Davey, J., Cornelissen, P. L., Thompson, H. E., Sonkusare, S., Hallam, G., Smallwood, J., & Jefferies, E. (2015). Automatic and controlled semantic retrieval: TMS reveals distinct contributions of posterior middle temporal gyrus and angular gyrus. *Journal of Neuroscience, 35*(46), 15230-15239.

Devereux, B. J., Clarke, A., Marouchos, A., & Tyler, L. K. (2013). Representational Similarity Analysis Reveals Commonalities and Differences in the Semantic Processing of Words and Objects. *Journal of Neuroscience, 33*(48), 18906-18916. doi:Doi 10.1523/Jneurosci.3809-13.2013

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27*(4), 429-446.

Fairhall, S. L., & Caramazza, A. (2013). Brain Regions That Represent Amodal Conceptual Knowledge. *Journal of Neuroscience, 33*(25), 10552-10558. doi:Doi 10.1523/Jneurosci.0051-13.2013

Gough, P. M., Nobre, A. C., & Devlin, J. T. (2005). Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. *Journal of Neuroscience, 25*(35), 8010-8016.

Graves, W. W., Binder, J. R., Desai, R. H., Conant, L. L., & Seidenberg, M. S. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *Neuroimage, 53*(2), 638-646. doi:10.1016/j.neuroimage.2010.06.055

Grindrod, C. M., Garnett, E. O., Malyutina, S., & den Ouden, D. B. (2014). Effects of representational distance between meanings on the neural correlates of semantic ambiguity. *Brain and Language, 139*, 23-35.

Halai, A. D., Parkes, L. M., & Welbourne, S. R. (2015). Dual-echo fMRI can detect activations in inferior temporal lobe during intelligible speech comprehension. *Neuroimage, 122*, 214-221.

Hebart, M. N., Görgen, K., & Haynes, J.-D. (2015). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Frontiers in neuroinformatics, 8*, 88.

Hendriks, M. H., Daniels, N., Pegado, F., & Op de Beeck, H. P. (2017). The effect of spatial smoothing on representational similarity in a simple motor paradigm. *Frontiers in neurology, 8*, 222.

Hino, Y., Lupker, S. J., & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology-Learning Memory and Cognition, 28*(4), 686-713.

Hoffman, P. (2019). Reductions in prefrontal activation predict off-topic utterances during speech production. *Nature communications, 10*(1), 515.

Hoffman, P., Binney, R. J., & Lambon Ralph, M. A. (2015). Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge. *Cortex, 63*, 250-265. doi:10.1016/j.cortex.2014.09.001

Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review, 125*, 293-328. doi:10.1037/rev0000094

Hoffman, P., & Woollams, A. M. (2015). Opposing effects of semantic diversity in lexical and semantic relatedness decisions. *Journal of Experimental Psychology: Human Perception and Performance, 41*, 385-402.

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 of the United States of America, 112*(25), 7857-7862.

Humphreys, G. F., & Lambon Ralph, M. A. (2014). Fusion and Fission of Cognitive Functions in the Human Parietal Cortex. *Cereb Cortex*. doi:10.1093/cercor/bhu198

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience, 18*, 665-679.

Jackson, R. L., Hoffman, P., Pobric, G., & Lambon Ralph, M. A. (2016). The Semantic Network at Work and Rest: Differential Connectivity of Anterior Temporal Lobe Subregions. *The Journal of Neuroscience, 36*(5), 1490-1501.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446.

Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex, 49*, 611-625.

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain, 129*(Pt 8), 2132-2147. doi:10.1093/brain/awl153

Jung, J., Cloutman, L. L., Binney, R. J., & Lambon Ralph, M. A. (2016). The structural connectivity of higher order association cortices reflects human functional brain networks. *Cortex*.

Krieger-Redwood, K., Teige, C., Davey, J., Hymers, M., & Jefferies, E. (2015). Conceptual control across modalities: graded specialisation for pictures and words in inferior frontal and posterior temporal cortex. *Neuropsychologia, 76*, 92-107.

Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience, 18*, 42-55. doi:10.1038/nrn.2016.150

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*, 211-240.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nat Rev Neurosci, 9*(12), 920-933. doi:10.1038/nrn2532

Makris, N., Goldstein, J. M., Kennedy, D., Hodge, S. M., Caviness, V. S., Faraone, S. V., . . . Seidman, L. J. (2006). Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophr Res, 83*(2), 155-171.

Mason, R. A., & Just, M. A. (2007). Lexical ambiguity in sentence comprehension. *Brain Research, 1146*, 115-127.

Mckiernan, K. A., Kaufman, J. N., Kucera-Thompson, J., & Binder, J. R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *Journal of Cognitive Neuroscience, 15*(3), 394-408.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mirman, D., Landrigan, J.-F., & Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological Bulletin, 143*(5), 499.

Musz, E., & Thompson-Schill, S. L. (2017). Tracking competition and cognitive control during language comprehension with multi-voxel pattern analysis. *Brain and Language, 165*, 21-32.

Nagel, I. E., Schumacher, E. H., Goebel, R., & D'Esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *Neuroimage, 43*(4), 801-807.

Noonan, K. A., Jefferies, E., Corbett, F., & Lambon Ralph, M. A. (2010). Elucidating the nature of deregulated semantic cognition in semantic aphasia: Evidence for the roles of the prefrontal and temporoparietal cortices. *Journal of Cognitive Neuroscience, 22*, 1597-1613.

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. *Journal of Cognitive Neuroscience, 25*(11), 1824-1850.

Ojemann, J. G., Akbudak, E., Snyder, A. Z., McKinstry, R. C., Raichle, M. E., & Conturo, T. E. (1997). Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *Neuroimage, 6*(3), 156-167. doi:DOI 10.1006/nimg.1997.0289

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

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.

Peelen, M. V., & Caramazza, A. (2012). Conceptual Object Representations in Human Anterior Temporal Cortex. *Journal of Neuroscience, 32*(45), 15728-15736. doi:Doi 10.1523/Jneurosci.1953-12.2012

Pexman, P. M., Hino, Y., & Lupker, S. J. (2004). Semantic ambiguity and the process of generating meaning from print. *Journal of Experimental Psychology-Learning Memory and Cognition, 30*(6), 1252-1270. doi:Doi 10.1037/0278-7393.30.6.1252

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.

Rice, G. E., Hoffman, P., & Lambon Ralph, M. A. (2015). Graded specialization within and between the anterior temporal lobes. *Annals of the New York Academy of Sciences, 1359*(1), 84-97.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making Sense of Semantic Ambiguity: Semantic Competition in Lexical Access. *Journal of Memory and Language, 46*(2), 245-266. doi:10.1006/jmla.2001.2810

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex, 15*(8), 1261-1269.

Rodd, J. M., Vitello, S., Woollams, A. M., & Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: An Activation Likelihood Estimation meta-analysis. *Brain Lang, 141*, 89-102. doi:10.1016/j.bandl.2014.11.012

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.

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Brecher, A., Faseyitan, O., Dell, G. S., . . . Coslett, H. B. (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 108*(20), 8520-8524.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14*(4), 489-537.

Shattuck, D. W., Mirza, M., Adisetiyo, V., Hojatkashani, C., Salamon, G., Narr, K. L., . . . Toga, A. W. (2008). Construction of a 3D probabilistic atlas of human cortical structures. *Neuroimage, 39*(3), 1064-1080.

Taylor, J. S. H., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin, 139*(4), 766-791.

Teige, C., Cornelissen, P. L., Mollo, G., Alam, T. R. d. J. G., McCarty, K., Smallwood, J., & Jefferies, E. (2019). Dissociations in semantic cognition: Oscillatory evidence for opposing effects of semantic control and type of semantic relation in anterior and posterior temporal cortex. *Cortex, 120*, 308-325.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 14792-14797.

Tyler, L. K., Cheung, T. P., Devereux, B. J., & Clarke, A. (2013). Syntactic computations in the language network: characterizing dynamic network properties using representational similarity analysis. *Frontiers in Psychology, 4*, 271.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*(6), 1176-1190.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience, 14*(4), 550-560.

Vitello, S., & Rodd, J. M. (2015). Resolving semantic ambiguities in sentences: Cognitive processes and brain mechanisms. *Language and Linguistics Compass, 9*(10), 391-405.

Westerlund, M., & Pylkkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia, 57*, 59-70.

Whitney, C., Grossman, M., & Kircher, T. T. (2009). The influence of multiple primes on bottom-up and top-down regulation during meaning retrieval: evidence for 2 distinct neural networks. *Cereb Cortex, 19*(11), 2548-2560. doi:10.1093/cercor/bhp007

Whitney, C., Jefferies, E., & Kircher, T. (2011). Heterogeneity of the left temporal lobe in semantic representation and control: priming multiple versus single meanings of ambiguous words. *Cereb Cortex, 21*(4), 831-844. doi:10.1093/cercor/bhq148

Xiang, H.-D., Fonteijn, H. M., Norris, D. G., & Hagoort, P. (2009). Topographical functional connectivity pattern in the perisylvian language networks. *Cerebral Cortex, 20*(3), 549-560.

Zempleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *Neuroimage, 34*(3), 1270-1279. doi:DOI 10.1016/j.neuroimage.2006.09.048

Supplementary Materials

Supplementary Table 1: Properties of primes in each condition

	Frequency	Concreteness	Length in letters	Length in syllables	Relatedness
Homonyms, Related	4.14 (0.65)	4.41 (0.62)	5.55 (1.88)	1.78 (0.83)	4.49 (0.51)
Homonyms, Unrelated	4.21 (0.62)	4.30 (0.71)	5.35 (1.93)	1.55 (0.60)	1.29 (0.20)
Unambiguous, Related	4.15 (0.57)	4.22 (0.84)	5.98 (1.62)	1.83 (0.75)	4.58 (0.37)
Unambiguous, Unrelated	4.01 (0.56)	4.33 (0.56)	6.05 (1.28)	1.65 (0.49)	1.22 (0.24)
Effect of Ambiguity	$F = 0.27$, $p = 0.61$	$F = 0.89$, $p = 0.35$	$F = 2.71$, $p = 0.10$	$F = 0.26$, $p = 0.61$	$F = 0.27$, $p = 0.61$
Effect of Relatedness	$F = 0.11$, $p = 0.74$	$F = 0.00$, $p = 0.98$	$F = 0.04$, $p = 0.85$	$F = 2.05$, $p = 0.15$	$F = 1938$, $p < 0.001$
Ambiguity x Relatedness interaction	$F = 0.75$, $p = 0.39$	$F = 0.69$, $p = 0.41$	$F = 0.17$, $p = 0.68$	$F = 0.03$, $p = 0.86$	$F = 1.19$, $p = 0.28$

Table reports mean values per condition (standard deviations in parentheses) and results of 2x2 ANOVAs contrasting conditions.
Frequency = Zipf frequency values from SUBTLEX corpus (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014).
Concreteness = values from (Brysbaert, Warriner, & Kuperman, 2014).
Relatedness = Ratings of semantic relatedness of prime & target on a scale of 1-5.

Supplementary Table 2: Peak activation co-ordinates for univariate contrasts of task

Effect	Location	Extent (mm^3)	*t*	MNI co-ords x	y	z
Semantic > Phonological	L ATL & IFG	33880				
	L anterior STS		9.54	-58	-2	-12
	L temporal pole/IFG		8.64	-44	20	-14
	L pars orbitalis		8.25	-50	34	-12
	R IFG & temporal pole	12584				
	R pars orbitalis		7.68	36	34	-14
	R temporal pole/IFG		6.11	44	16	-18
	R anterior STS		5.49	52	-4	-22
	R cerebellum	4648				
	R cerebellum		6.45	10	-84	-28
	R cerebellum		6.38	12	-80	-36
	R cerebellum		3.58	18	-92	-42
	L inferior parietal lobe	8672				
	L intraparietal sulcus		6.22	-32	-78	48
	L angular gyrus		6.04	-46	-78	26
	L temporoparietal junction		4.74	-56	-66	24
	Ventromedial PFC	5784				
	Gyrus rectus		5.90	-4	50	-14
	Gyrus rectus		5.26	-6	36	-18
	Subgenual cingulate		4.20	-2	20	-12
	Dorsomedial & L dorsolateral PFC	22952				
	Pre-supplementary motor area		5.76	-4	30	56
	Superior frontal gyrus		5.60	0	42	42
	Pre-supplementary motor area		5.25	-4	28	66
Phonological > Semantic	L inferior & superior parietal lobe	34824				
	L intraparietal sulcus		10.87	-48	-40	46
	L superior parietal lobule		6.68	-22	-64	52
	L intraparietal sulcus		6.07	-44	-48	62
	L dorsolateral PFC	8304				
	L middle frontal gyrus		7.30	-44	40	20
	L middle frontal gyrus		3.56	-30	34	32
	L middle frontal gyrus		3.29	-30	48	32
	Medial PFC & L precentral gyrus	24320				
	Supplementary motor area		7.29	-4	4	68

Precentral gyrus		6.92	-2	-18	64
L superior frontal gyrus		6.65	-22	2	52
R inferior parietal lobe	18064				
R postcentral gyrus		6.83	48	-30	52
R intraparietal sulcus		5.29	42	-44	56
R occipital cortex		3.98	28	-64	40
R cerebellum	5560				
R cerebellum		5.59	28	-62	-52
R cerebellum		3.97	12	-78	-52
R cerebellum		3.33	32	-46	-50
R occipitotemporal & cerebellum	5664				
R posterior fusiform gyrus		5.35	30	-68	-22
R lingual gyrus		4.83	8	-58	-6
R lingual gyrus		3.97	12	-60	4
R dorsolateral PFC	5544				
R middle frontal gyrus		5.03	42	42	16
R middle frontal gyrus		3.47	36	40	26
R frontal white matter		3.39	24	40	24

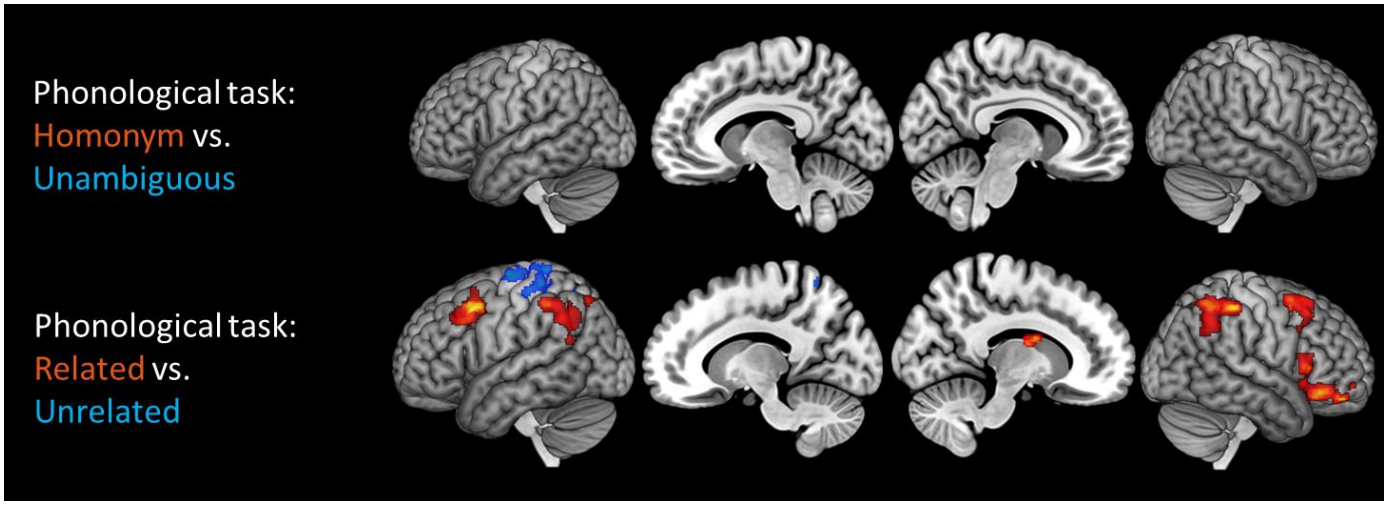
ATL = anterior temporal lobe; IFG = inferior frontal gyrus; STS = superior temporal sulcus; PFC = prefrontal cortex.

Supplementary Table 3: Peak activation co-ordinates for univariate contrasts in the meaning task

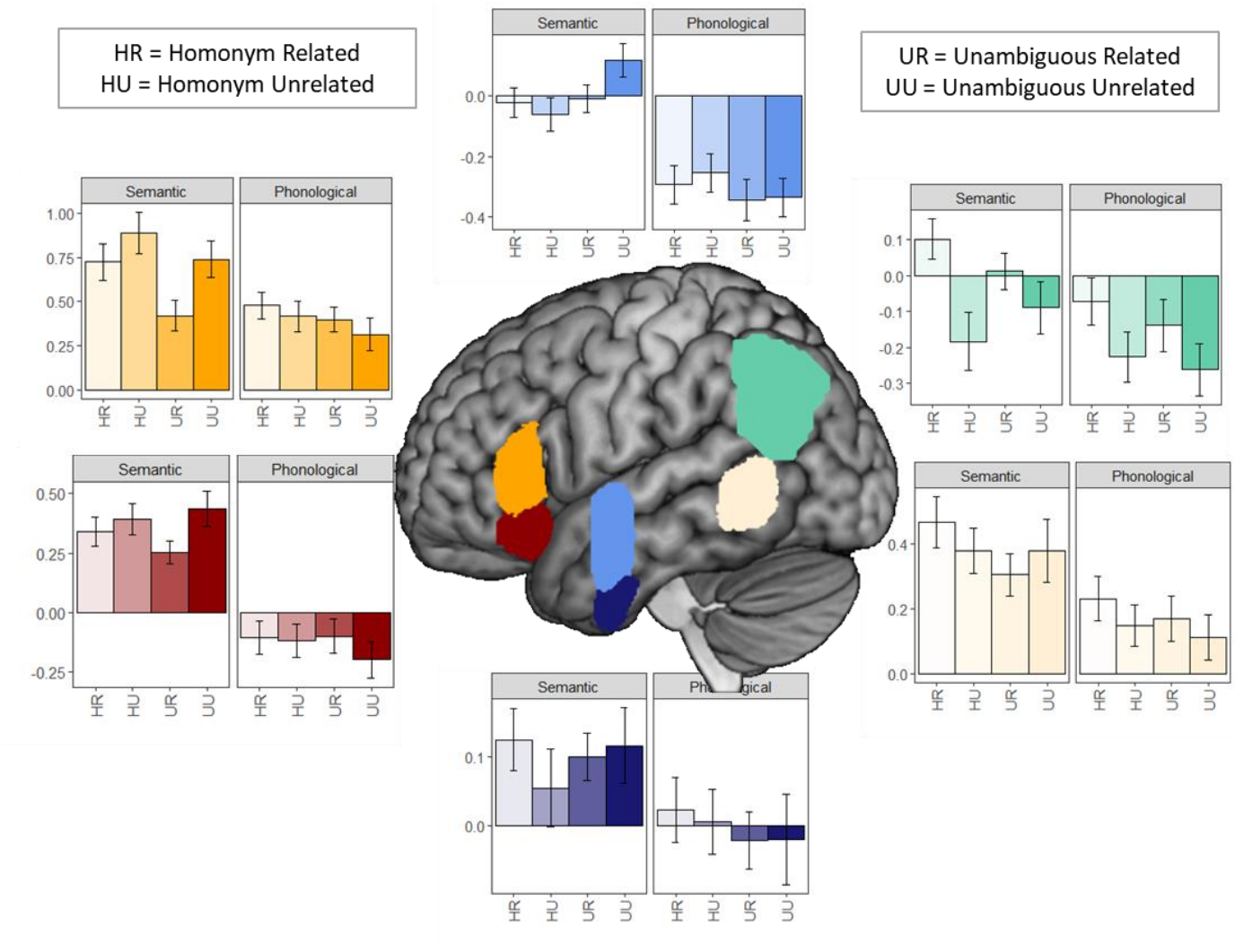
Effect	Location	Extent (mm^3)	*t*	MNI co-ords x	y	z
Homonym > Unambiguous	L lateral PFC	23384				
	L pars triangularis		6.56	-40	32	14
	L pars triangularis		6.01	-38	14	30
	L pars opercularis		5.90	-34	8	24
	L intraparietal sulcus	7408				
	L intraparietal sulcus		5.20	-30	-58	48
	L middle occipital gyrus		4.97	-30	-70	42
	L intraparietal sulcus		4.93	-44	-44	52
	Bilateral cerebellum	4496				
	R cerebellum		5.13	26	-72	-50
	L cerebellum		4.33	-16	-66	-28
	L cerebellum		4.10	16	-76	-46
	Dorsomedial PFC	4856				
	Presupplementary motor area		4.94	-8	22	48
	Presupplementary motor area		3.96	-2	32	48
	L anterior cingulate		3.76	-12	28	34
	L posterior temporal	7296				
	L ITG		4.69	-46	-52	-12
	L ITG/MTG		4.14	-52	-68	-6
	L ITG		4.11	-56	-62	-14
Unambiguous > Homonym	Rostromedial PFC	5456				
	L medial PFC		4.43	-6	56	0
	Frontal pole		4.22	-4	58	10
	Gyrus rectus		3.82	0	52	-10
Related > Unrelated	R inferior parietal & occipital	15096				
	R supramarginal gyrus		7.74	60	-32	54
	R angular gyrus		4.79	50	-46	42
	R supramarginal gyrus		4.63	50	-38	32
	L inferior parietal lobe	19776				
	L supramarginal gyrus		6.56	-56	-42	38
	L angular gyrus		6.42	-62	-50	42
	L angular gyrus		5.81	-52	-64	38

	Ventromedial PFC	10832				
	Paracingulate gyrus		6.42	0	48	10
	L frontal pole		5.23	-12	58	4
	Paracingulate gyrus		5.23	8	48	12
	L mid-lateral temporal lobe	5000				
	L middle temporal gyrus		5.74	-66	-34	-8
	L middle temporal gyrus		5.29	-58	-26	-18
	Posterior cingulate	6008				
	Precuneus		4.78	-2	-44	48
	Posterior cingulate cortex		4.64	-4	-18	38
	Posterior cingulate cortex		4.30	0	-30	44
Unrelated > Related	L lateral PFC	13728				
	L pars triangularis		7.24	-42	28	10
	L precentral gyrus		6.47	-40	-2	34
	L middle frontal gyrus		5.94	-56	30	26
	Dorsomedial PFC	5848				
	R anterior cingulate		5.99	10	22	42
	L anterior cingulate		5.94	-10	14	46
	L posterior ventral temporal	4856				
	L fusiform gyrus		5.15	-40	-36	-20
	L inferior temporal gyrus		4.69	-46	-48	-14
	L fusiform gyrus		3.90	-30	-34	-22

PFC = prefrontal cortex; MTG = middle temporal gyrus; ITG = inferior temporal gyrus.



<u>Supplementary Figure 1: Whole-brain univariate activation contrasts for the phonological task</u> Images are shown at a threshold of p<0.05, corrected for multiple comparisons at the cluster level.



<u>Supplementary Figure 2: Activation in each region of interest as a function of condition</u> Effect size is shown relative to implicit resting baseline. Bars indicate 1 SEM.