
Article:

Teige, Catarina Ingrid R, Mollo, Giovanna, Millman, Rebecca E. et al. (4 more authors) (2018) Dynamic semantic cognition : Characterising coherent and controlled conceptual retrieval through time using magnetoencephalography and chronometric transcranial magnetic stimulation. Cortex; a journal devoted to the study of the nervous system and behavior. pp. 329-349. ISSN 1973-8102

https://doi.org/10.1016/j.cortex.2018.03.024

Dynamic semantic cognition: Characterising coherent and controlled conceptual retrieval through time using magnetoencephalography and chronometric transcranial magnetic stimulation

Catarina Teige[1], Giovanna Mollo[1], Rebecca Millman[2], Nicola Savill[1], Jonathan Smallwood[1], Piers L. Cornelissen*[3], Elizabeth Jefferies*[1]

[1]Department of Psychology and York Neuroimaging Centre, University of York

[2]Manchester Centre for Audiology and Deafness, Division of Human Communication, Development and Hearing, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester

[3]Department of Psychology, Northumbria University

*Correspondence to: Piers Cornelissen, Psychology Department, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK

OR

Elizabeth Jefferies, Department of Psychology, University of York, York, YO10 5DD, UK

Abstract

Distinct neural processes are thought to support the retrieval of semantic information that is (i) coherent with strongly-encoded aspects of knowledge, and (ii) non-dominant yet relevant for the current task or context. While the brain regions that support readily coherent and more controlled patterns of semantic retrieval are relatively well-characterised, the temporal dynamics of these processes are not well-understood. This study used magnetoencephalography (MEG) and dual-pulse chronometric transcranial magnetic stimulation (cTMS) in two separate experiments to examine temporal dynamics during the retrieval of strong and weak associations. MEG results revealed a dissociation within left temporal cortex: anterior temporal lobe (ATL) showed greater oscillatory response for strong than weak associations, while posterior middle temporal gyrus (pMTG) showed the reverse pattern. Left inferior frontal gyrus (IFG), a site associated with semantic control and retrieval, showed both patterns at different time points. In the cTMS experiment, stimulation of ATL at ~150ms disrupted the efficient retrieval of strong associations, indicating a necessary role for ATL in coherent conceptual activations. Stimulation of pMTG at the onset of the second word disrupted the retrieval of weak associations, suggesting this site may maintain information about semantic context from the first word, allowing efficient engagement of semantic control. Together these studies provide converging evidence for a functional dissociation within the temporal lobe, across both tasks and time.

Keywords: comprehension; beamforming; semantic memory; anterior temporal; posterior middle temporal gyrus

Semantic cognition allows us to understand the meaning of our environment to drive appropriate thoughts and behaviour. It comprises several distinct yet interacting components (Jefferies, 2013; Jefferies & Lambon Ralph, 2006; Lambon Ralph, Jefferies, Patterson & Rogers, 2017). Semantic representations capture the meanings of words and objects across contexts, supporting coherent conceptual retrieval from fragmentary inputs and generalisation across situations. However, the retrieval of specific aspects of knowledge in a context-dependent fashion requires control mechanisms that shape evolving retrieval towards currently-pertinent semantic features, and away from dominant yet irrelevant associations. While patterns of activation within the semantic store may be sufficient to uncover links between items that are highly coherent with long-term knowledge (i.e. items that share multiple features or are frequently associated, such as *pear-apple* or *tree-apple*), additional control processes may be required to recover non-dominant aspects of knowledge, such as *worm-apple*, since strong but currently-irrelevant associations (e.g. *worm-soil*) must be disregarded (Lambon Ralph, Jefferies, Patterson & Rogers, 2017; Gold et al., 2006).

Although the brain regions that support semantic cognition are relatively well-described, the temporal dynamics are not. Neuroimaging studies have highlighted the importance of a distributed left-dominant network underpinning semantic cognition, including anterior temporal lobe (ATL), posterior middle temporal gyrus (pMTG) and inferior frontal gyrus (IFG) (Jefferies, 2013; Vandenberghe et al., 1996; Xu, Qixiang, Zaizhu, Yong, & Yanchao, 2016; Lambon Ralph et al., 2017; Binder, Desai, Graves & Conant, 2009). These brain regions make dissociable contributions to semantic cognition, although their specific roles remain controversial. The ventral ATL is proposed to support heteromodal concepts that are extracted from multiple inputs (e.g., vision, audition, smell; Patterson, Nestor & Rogers, 2007; Lambon Ralph et al., 2017). Patients with semantic dementia (SD), show progressive degradation of knowledge across modalities following atrophy and hypometabolism in ATL (Mion et al., 2010; Bozeat et al., 2000; Rogers et al., 2006). Convergent evidence for a role of ATL in multimodal conceptual processing is provided by positron emission tomography (e.g. Bright et al., 2004; Crinion et al., 2003; Devlin et al., 2000; Noppeney & Price, 2002; Rogers et al., 2006; Scott et al., 2000), functional magnetic resonance imaging (fMRI) – particularly when magnetic susceptibility artefacts within ATL are minimised (Binney et al., 2010; Murphy et al., 2017; Visser et al.,2010; 2012), magnetoencephalography (MEG) (Lau et al, 2013; Clarke et al., 2011; Marinković et al., 2003; Fujimaki et al., 2009; Mollo et al., 2017), intracranial electrode arrays (Chan et al., 2011; Chen et al., 2016) and transcranial magnetic stimulation (TMS) (Lambon Ralph et al., 2009; Pobric et al., 2007; 2009). The ATL is allied to the default mode network (DMN) in terms of connectivity and function (Binder et al., 2003; Davey et al., 2015; 2016; Wirth et al., 2011), although the maximal semantic response in ATL is not identical to the site of peak DMN connectivity (Jackson et al., 2016). In common with DMN regions, ATL shows a larger response to easy or more automatic aspects of semantic retrieval, such as identifying dominant aspects of knowledge (e.g., linking DOG with CAT; Davey et al., 2016), and when coherent meaning emerges from conceptual combinations (Bemis & Pylkkänen, 2012; Hoffman, Binney & Lambon Ralph, 2015). ATL is also implicated in semantic retrieval during mind-wandering (Binder et al., 1999; Smallwood et al., 2016). Collectively, these findings suggest that ATL responds most strongly when ongoing semantic retrieval is highly coherent with long-term knowledge – although *causal* evidence is currently lacking.

Brain regions distinct from ATL are implicated in the *control* of semantic cognition. The contribution of left IFG to executive-semantic processes has been documented by many fMRI studies (e.g., Thompson-Schill, D'Esposito, Aguirre &Farah, 1997; Badre, Poldrack, Pare-Blagoev, Insler & Wagner, 2005; Noppeney, Phillips & Price, 2004; Noonan et al., 2013; Bedny, McGill & Thompson-Schill, 2008). Convergent evidence for a causal contribution of this region has been provided by transcranial magnetic stimulation (TMS, Hoffman, Jefferies & Lambon Ralph, 2010; Whitney et al., 2011) and neuropsychology: patients with damage to left IFG have difficulty flexibly tailoring their semantic retrieval to suit the circumstances (Thompson-Schill et al., 1998; Jefferies & Lambon Ralph, 2006; Corbett, Jefferies & Lambon Ralph, 2009; Thompson et al., 2015). While the contributions of ATL and IFG align with recent component process views of semantic cognition (e.g., the Controlled Semantic Cognition framework, which suggests semantic cognition reflects an interaction of conceptual representations and control processes, Lambon Ralph et al., 2017), the contribution of pMTG remains controversial. Some accounts have proposed that posterior temporal areas provide an important store of conceptual representations (Martin, 2007), with pMTG specifically implicated in knowledge of actions and events (Chao, Haxby & Martin, 1999; Martin et al., 1995). Alternatively, a growing literature supports the view that pMTG is part of a distributed network with IFG and other regions underpinning semantic control (Vitello et al., 2014; Jefferies, 2013; Davey et al., 2016; Noonan et al., 2013; Gold et al., 2006). A meta-analysis showed that a widely distributed set of cortical regions is reliably activated across diverse manipulations of semantic control demands, with left pMTG showing the second most consistent response after left IFG (Noonan et al., 2013). Semantic control deficits can follow from either left prefrontal or posterior temporal lesions (Jefferies & Lambon Ralph, 2006; Noonan et al., 2010); moreover, inhibitory TMS to left pMTG and IFG produces equivalent disruption of semantic judgements that require controlled but not automatic retrieval (Whitney et al., 2011; Davey et al., 2015), and inhibitory stimulation of IFG elicits a compensatory increase in pMTG (Hallam et al., 2017). These regions also show a strong pattern of both structural and functional connectivity (Davey et al., 2016; Hallam et al., 2016; Jeyoung & Lambon Ralph, 2016), consistent with the view that they form a large-scale distributed network underpinning controlled aspects of semantic retrieval, and they respond to semantic control demands across modalities (Krieger-Redwood et al., 2015). This semantic control network partially overlaps with the frontoparietal control network and thus both left IFG and pMTG have a different pattern of large-scale connectivity from ATL (Davey et al., 2016).

Component process accounts of semantic cognition (e.g. Jefferies, 2013; Lambon Ralph et al., 2017) predict a functional dissociation within the temporal lobe – with ATL supporting efficient retrieval when currently-relevant semantic information is highly coherent with dominant aspects of long term knowledge, and pMTG (along with IFG) playing a critical role at times when such knowledge cannot serve the goal of the moment. The current work tests this predicted dissociation by examining how the contribution of these sites changes when dominant conceptual associations no longer support appropriate patterns of retrieval, and more weakly-encoded information is required. We presented two words successively and manipulated the strength of the relationship between them. When two words are strongly associated, retrieval of the relevant conceptual link is thought to be relatively automatic, since the meaning of the second word strongly overlaps with features activated from the first word. Consequently, the first word establishes a pattern of semantic activation that is strongly coherent with emerging conceptual retrieval to the second word without additional constraints. For weaker associations, semantic activation elicited from the first word is less coherent with the second input, and with the pattern of conceptual retrieval that is needed to elicit the correct response. Consequently, semantic control processes are thought to be engaged to shape conceptual retrieval so that it is relevant to the current context (cf. Badre et al., 2005; Whitney et al., 2011). This contrast of weak and strong associations has been used commonly in previous fMRI and TMS investigations of semantic control (e.g., Badre et al., 2005; Whitney et al., 2011; Davey et al., 2016). The associated words were presented successively to provide a clear temporal marker (the onset of the second word) from which to examine more coherent and controlled patterns of semantic retrieval, and both strong and weakly-associated word pairs required the same response (to indicate that ‘yes’ the words were related).

If different sites play distinct roles in automatic and controlled semantic retrieval, we reasoned that, as well as overall differences in their response to strong and weak associations, there might also be differences in the timing of these effects. Little is known about differences in the time-course of semantic retrieval at different sites – and previous work has often used electroencephalography (EEG), which may lack the spatial resolution to separately resolve signals from ATL and pMTG. MEG studies of ATL show early responses (from 120ms) that appear to reflect interactions between semantic representations and inputs (Clarke et al., 2011; Mollo et al., 2017), plus later responses (250-450ms) that are influenced by patterns of coherent conceptual retrieval across both modalities (Marinkovic et al., 2003) and multiple items (Halgren et al., 2002; Lau et al., 2013; Bemis & Pylkkänen, 2011). Moreover, a recent chronometric TMS study by Jackson et al. (2015)

found that the critical time point of involvement for ATL was around 400ms (although this study did not manipulate the strength of association and thus cannot identify when semantic processing in ATL is critical for the efficient retrieval of more coherent concepts). An N400 response has also been localised to pMTG (Helenius, Salmelin, Service & Connolly, 1998; Halgren et al., 2002; Lau, Phillips & Poeppel, 2008). This N400 effect is greater for unexpected meanings (Brown & Hagoort, 1993; Maess et al., 2006), although it also responds to a wide variety of semantic and lexical manipulations (Halgren et al., 2002; Lau, Phillips & Poeppel, 2008). In line with the N400 literature, ATL and pMTG can show a similar response to violations of semantic expectations – i.e., a stronger response to unrelated than related items (for a review, see Lau et al., 2008) – and thus the N400 semantic priming effect does not readily distinguish between ATL and pMTG; however, research has linked ATL to relatively automatic semantic priming (Lau et al., 2013) and the response in pMTG to more controlled or strategic semantic priming (Gold et al., 2006). E/MEG work has shown that the response to unexpected meanings corresponds to a decrease in oscillatory power in the beta band, suggesting that oscillatory activity in this frequency range might support the maintenance of an appropriate network for comprehension given current expectations (Luoa et al., 2010; Wang et al., 2012; Kielar et al., 2014; Lewis & Bastiaansen, 2015). A recent MEG study also identified a potential role for both left IFG and pMTG in the detection and implementation of a suitable context for semantic retrieval (indicated by presentation of the words 'to' or 'the', which established a noun or verb-based interpretation); this response was stronger for ambiguous words within the first 100ms after presentation, suggesting that semantic control sites might play a role in initiating control processes when these are required, even before the meaning of a word has been fully accessed, as well as selecting specific interpretations or features at a later stage (Mollo et al., 2018 in press; see also Ihara et al., 2007).

Here, we used two temporally-sensitive methods (MEG, chronometric TMS) to examine the engagement of left ATL and pMTG in semantic retrieval through time. In the MEG experiment, we also characterised responses within left IFG. By manipulating the strength of association between two words during explicit semantic decisions, we were able to test predictions of the Controlled Semantic Cognition framework (Lambon Ralph et al., 2017; Jefferies, 2013). By this account, left IFG and pMTG are expected to show a stronger response to weak compared with strong associations, consistent with a role in controlled aspects of semantic retrieval, while ATL is predicted to show a stronger response to items coherent with dominant aspects of knowledge (i.e., effects of strong > weak associations). We were also able to test two alternative hypotheses about the timing of these effects. By one view, information is first retrieved and then selected: this account might envisage effects of strong associations in ATL that precede the engagement of controlled retrieval in pMTG and IFG. Alternatively, controlled retrieval processes may be engaged at an early stage following the onset of the second word, to shape patterns of semantic retrieval in ATL. For example, semantic control regions might maintain a semantic ‘context’ (or meaning-based biases) that can shape retrieval such that it suits the circumstances: by this view, ATL maintains long-term conceptual relationships, while the information represented in pMTG and/or IFG might change rapidly and adaptively, reflecting recent and currently-relevant conceptual information. An emerging lack of coherence between the processing of a new input and the existing semantic context might be critical to the engagement of semantic control processes in pMTG and IFG. Consequently, these regions might be engaged rapidly following the onset of a weakly-associated second word. These predictions are motivated by emerging evidence that visual word processing does not occur in a linear sequence from orthography to semantic and/or articulatory representations (Klein et al., 2012; Sereno et al., 2003; Wheat et al., 2010; Woodhead et al., 2014; Yvert et al., 2012). Instead, the interplay between vision and meaning will depend on the extent to which input processing is supported by the current state of conceptual representations; effects of associative strength could then emerge in advance of full semantic retrieval to the second item in a sequence.

In Experiment 1, we used beamforming analyses to characterise changes in total oscillatory power in ATL, pMTG and IFG during the retrieval of strong and weak associations. Total power includes components that are not phase-locked to an event/stimulus (i.e., responses that are generated at a slightly different time point across trials or participants). These so-called “induced” responses might be prominent in the retrieval of semantic relationships that span successive items (since the emergence of relationships between inputs might not be time and phase-locked to the onset of the second word). In Experiment 2, chronometric TMS was used to determine the causal role that anterior and posterior regions of the temporal lobe play in the retrieval of strong and weak associations at different time points. Together these two experiments, using different neuroscientific techniques, characterise the spatiotemporal basis of semantic retrieval when information is coherent with strongly-encoded aspects of knowledge, and show how this changes when non-dominant aspects of knowledge are required.

Experiment 1: MEG

Materials and Methods:

Participants:

Participants were 20 right-handed native English speakers, with normal or corrected-to-normal vision, and no history of language disorders (14 female, mean age 23.3 years, range 20-35). Data from one participant was excluded because their accuracy in the task fell below the acceptable minimum of 75% correct. Written consent was obtained from all participants and the study was approved by the York Neuroimaging Centre Research Ethics Committee.

Materials:

The task and stimuli were adapted from Badre et al. (2005). Word pairs were presented, one word at a time, with varying associative strength between the first and second word, and participants were asked to decide if the two words were related in meaning or not. Participants were presented with 440 word pairs that were strongly-related (n=110), weakly-related (n=110), or unrelated (n=220). The correct response was the same for strong and weak trials (i.e., participants had to indicate that these words were related). Strong and weak word pairs were selected using free association response data from the Edinburgh Associative Thesaurus (EAT). Strong associates were produced relatively frequently by participants (23%), while weak associates were produced more rarely (1%). The difference in mean association strength between strong and weak conditions was highly significant (t(188)=16.05, *p*<.001; Table 1). The first word was expected to initiate a pattern of conceptual retrieval (i.e., biases within the semantic system); semantic retrieval to the second word then occurred within this context. The biases established by the first word facilitated the semantic decision substantially for strongly-associated trials and rather less for weakly-associated trials.

While our analyses focussed on the second word in each pair (which were identical across conditions between subjects), Table 1 confirms that there were no significant differences in word frequency or length across strong and weak conditions for the initial word. While the words were related in diverse ways, there were also no significant differences between the weak and strong conditions in the frequency of these different kinds of semantic association. The associations could be: (i) categorical (e.g., deer-cow: 40% and 37% of strong and weak trials respectively); (ii) thematic (e.g., soup-bowl: 42% and 50% of strong and weak trials); (iii) both categorical and thematic (e.g., holly-ivy: 8% and 4% of trials); (iv) part-whole relationship (e.g., pony-mane: 10% and 8% of trials) and (v) linguistic-only (e.g., pop-corn: <1%). A chi-square analysis including the four most frequent trial types revealed no significant difference between the strong and weak conditions (χ^2 = 4.4, p = .22). While very few trials fell within the linguistic-only category, most of the semantically-related items were also linguistically related and we are not able to separate the effects of these factors within the current study (although as noted in the Introduction, both ATL and pMTG are associated with semantic processing across modalities).

Table 1: Comparing word frequency and length for the first word across conditions, plus the associative strength between the two words in the MEG experiment

Measure	Strong Association M (SD)	Weak Association M (SD)	p-value
Word frequency	26.6 (64.20)	29.1 (38.0)	.59
Word length (letters)	5.5 (1.80)	5.0 (1.5)	.16
Association strength	0.23 (0.19)	0.01 (0.005)	.001

Unrelated trials were created by randomly shuffling words across pairs and manually removing any associations arising by chance. Target words were presented *either* following a strong or weak associate (not both), and in the unrelated condition. This meant that there was a 50% chance on any trial that a pair of words was semantically related.

Procedure:

An illustration of the procedure can be seen in Figure 1a. Nonius lines (acting as a fixation cross) were present at all times. Before each trial, there was a rest period of 800 ms, plus an unpredictable jittered interval from 0 to 1000 ms, designed to reduce anticipatory responses. The first word was presented for 200 ms, there was an inter-stimulus interval (ISI) of 150 ms, and then the second word appeared for 200 ms followed by a 1000 ms interval. After each trial, the nonius lines were dimmed (for 1200 ms) and participants were encouraged to confine blinking to this period. The task required participants to make an explicit judgement about the relationship between the two words. On 10% of the trials, participants were cued to make an overt response by the presence of a question mark (on screen for 1000ms). They pressed one of two buttons with their left hand to indicate whether they had identified an association. These ‘catch trials’ were used to monitor performance in the task, and were excluded from further analysis. Since we only collected behavioural data for a small number of trials during MEG (to keep participants attending to the task), we also ran a behavioural version of the experiment outside the scanner, with the same participants, a minimum of 4 weeks before MEG data collection. This experiment was identical to the MEG version, except a response was given on every trial, and the pairings between stimuli were reversed – if a particular target was paired with a strong associate in the behavioural experiment, it was presented following a weak associate in MEG (and vice versa). Data from the behavioural experiment and the catch-trials in MEG are shown in Figure 1b and c.

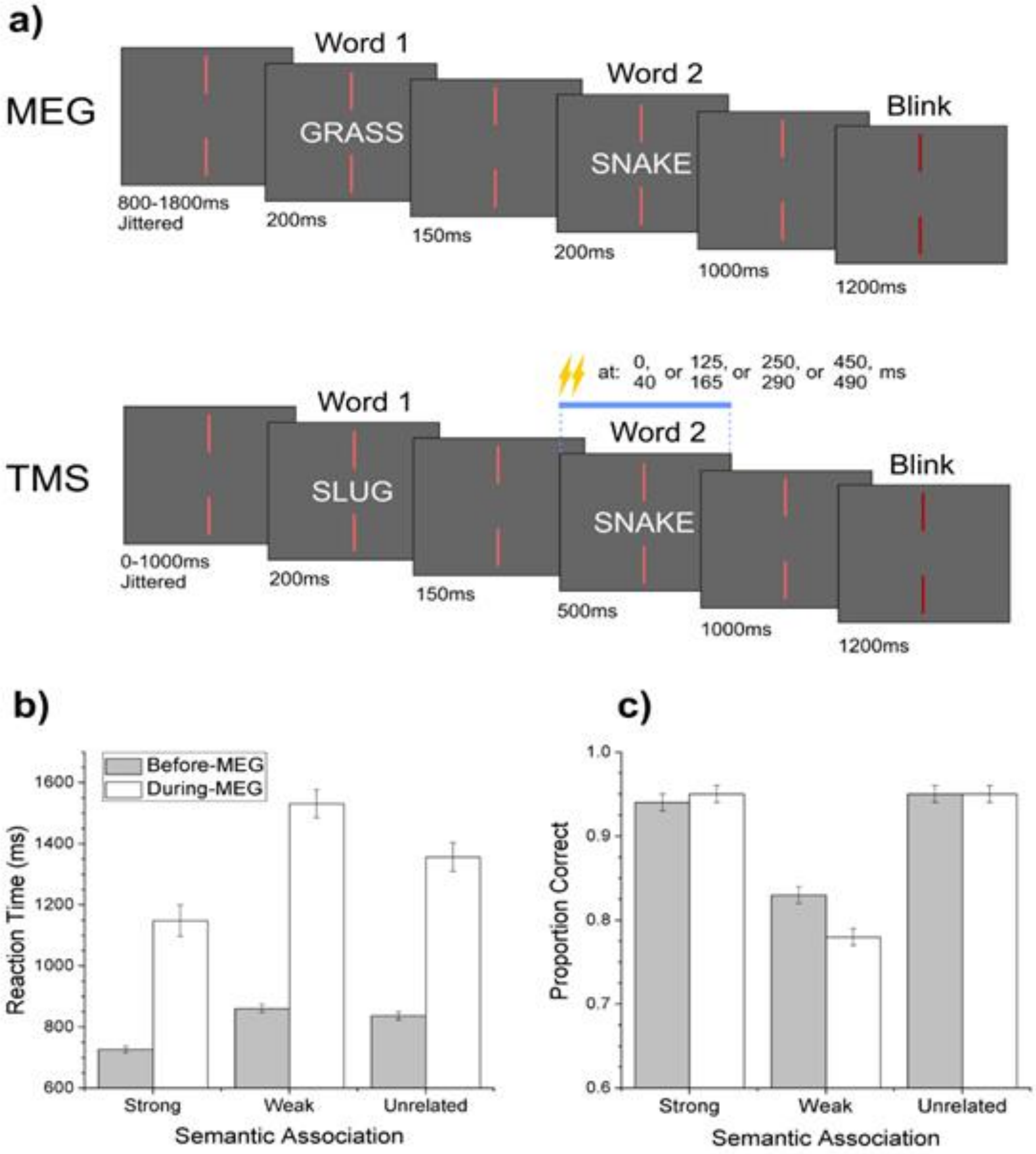


Figure 1: a) Example trials and timelines for the MEG and TMS experiments. b) Reaction time and c) accuracy data from the behavioural (gray bars) and MEG (white bars) experiments. Standard errors are corrected for repeated measures.

Stimulus presentation:

The MEG experiment was carried out in a dark, magnetically shielded room. Presentation version 16.1 (Neurobehavioral Systems) was used to present the stimuli, communicate their timings to the MEG data acquisition system and to record participants' responses on catch trials. Stimuli were back-projected onto a screen with a viewing distance of ~75 cm, so that letter strings subtended ~1˚ vertically and ~5˚ horizontally at the retina. Light grey letters on a dark grey background were used, such that the screen luminance was in the mesopic range, and a neutral density filter was used to minimize glare.

Data collection:

During MEG recordings, participants were seated in an upright position, with the magnetometers arranged in a helmet shaped array, using a whole-head 248-channel, Magnes 3600 system (4D Neuroimaging, San Diego, California). Data were recorded in continuous mode, with a sampling rate of 678.17 Hz and pass-band filtered between 1-200 Hz. Electrooculography was not recorded.

Before MEG data acquisition, participants' head shape and the location of five head coils were recorded with a 3D digitizer (Fastrak Polhemus). The signal from the head coils was used to localize the participant's head position with respect to the magnetometer array before and after the experiment. The 3D digitized head shape of each participant was used for the co-registration of individual MEG data onto the participant's structural MRI image using a surface-based alignment procedure from Kozinska, Carducci, and Nowinski (2001). For each participant, a high-resolution structural T1-weighted anatomical volume was acquired in a GE 3.0 T Signa Excite HDx system (General Electric, USA) at the York Neuroimaging Centre, University of York, with an 8-channel head coil and a sagittal-isotropic 3-D fast spoiled gradient-recalled sequence (repetition time/echo time/flip angle = 8.03 msec/3.07 msec/20°, spatial resolution of 1.13 mm × 1.13 mm × 1.0 mm, in-plane resolution of 256 × 256 × 176 contiguous slices).

External, non-biological noise detected by the MEG reference channels was subtracted, and MEG data were converted into epochs of 1500 ms length, starting 800 ms before the onset of the first word. All channels from all trials were inspected visually and epochs containing artifacts, such as blinks, articulatory movements, swallowing, and other movements, were rejected manually. Data from three malfunctioning channels were automatically rejected for all participants. Statistical analyses included only datasets with at least 75% of trials retained after artefact rejection. 20 (of 21) datasets reached this criterion. On average, 17% of the trials were rejected from these datasets (min 7.3% - max 25%).

MEG analysis strategy:

Our analysis strategy involved first localising effects for the paradigm across the entire brain, collapsing across conditions, and then selecting points of interest that showed a strong response to the task as a whole and that also related to hypothesised sites relevant to semantic control from the literature (sites which were additionally used as the focus for TMS stimulation). This analysis approach allows us to characterise differences between conditions in time and frequency within sites of interest relevant to the Controlled Semantic Cognition hypothesis (Jefferies, 2013; Lambon Ralph et al., 2017), without biasing the MEG analysis towards the hypothesised differences between conditions. In previous studies using a similar approach, condition differences in spectral plots were often localised to specific frequencies and points in time, as opposed to reflecting differences in the mean signal (e.g., Klein et al., 2014, Cerebral Cortex). Consequently, we expected the whole-brain beamforming that constituted the first step of our analysis to be largely insensitive to condition differences: these analyses necessarily aggregate data across relatively wide time windows and frequency bands in order to provide full coverage of the parameter space. Given our reliance on ROIs, the study was optimised to characterise the nature of differences at points-of-interest, as opposed to localising these effects across the whole brain. This approach adds value to the literature since previous fMRI studies have already localised regions that respond to relatively automatic and more controlled patterns of semantic retrieval (e.g., Davey et al., 2016; Noonan et al., 2013), yet the temporal evolution of semantic retrieval at these sites is less clear.

For both whole-brain and POI analyses, sources of neural activity were reconstructed with a modified version of the vectorised, linearly-constrained minimum-variance (LCMV) beamformer (Van Veen et al., 1997; Huang et al., 2004), implemented as part of the public domain Neuroimaging Analysis Framework (NAF) pipeline at the York Neuroimaging Centre (http://vcs.ynic.york.ac.uk/docs/naf/index.html). We used a multiple spheres head model (Huang et al., 1999), with co-registration checked manually. An MEG beamformer (spatial filter) allows the signal coming from a location of interest in the brain to be estimated while attenuating signals from elsewhere. This is achieved by reconstructing the neuronal signal at a specific point (referred to as a Virtual Electrode) as the weighted sum of the signals recorded by the MEG sensors. The covariance matrix, used to generate the weights for each beamformer, was regularized using an estimate of noise covariance (Prendergast et al., 2011; Hymers et al., 2010). This procedure was performed separately for each condition and/or analysis window, in order to optimise sensitivity to the effect of interest (Brookes et al., 2008; 2011). The outputs of the three spatial filters at each point in the brain were summed to generate estimates of oscillatory power. This analysis strategy and the parameters used for the current study were similar to those used in recent MEG studies of visual word recognition, object naming and semantic processing (Wheat et al., 2010; Klein et al., 2014; Urooj, 2014; Mollo et al., 2017).

Whole brain beamforming:

The brain's overall response to the task (collapsing the strong and weak trials) was characterised within broad frequency ranges and relatively long periods of 200ms. A cubic lattice of points was defined in the brain (5 mm spacing), and at each point, an independent set of spatial filters was defined to estimate the source current at that point. A noise-normalised volumetric map of total oscillatory power (i.e., including both the evoked and non-phase locked components) was then produced over these broad temporal windows and frequency bands. Since our main research question concerned how the brain's response to the second word changed as a function of its relationship to the first word, we defined time zero as the onset of the second word of the pair; the onset of the first word was at -350ms relative to this. We then characterised whole-brain oscillatory responses to the second word by contrasting responses in "active" time windows at 0-200ms, 200-400ms, and 400-600ms post-onset of the second word with a 200ms "passive" time window at -700 to -500ms (prior to the onset of the trial). The Neural Activity Index (NAI; Van Veen et al., 1997), which is an estimate of oscillatory power that takes account of spatially-inhomogeneous noise, was calculated at each point in the lattice, within the following frequency pass-bands: 5-15 Hz, 15-25 Hz, 25-35 Hz and 35-50 Hz. These frequency ranges were taken from previous MEG studies of reading (Klein et al., 2014; Wheat et al., 2014). This analysis produced an NAI volumetric map for the active and passive period, separately for each participant at each frequency band, from which paired-samples t-statistics were calculated. Individual participant's t-maps were then transformed into the MNI standardized space in order for group level statistics to be calculated. To do this, a null distribution was built up by randomly relabelling the active and passive windows for each participant at each grid point, using the permutation procedure developed by Holmes et al. (1996). The maximum t-value obtained with random relabelling across 10000 permutations was established. We then compared the real distribution of t-values in the data with the maximum t-value obtained from the permuted data. Maximum statistics can be used to overcome the issue of multiple comparisons (i.e. controlling experiment-wise type I error), since the approach uses the highest permuted t-value across the brain to provide a statistical threshold for the whole lattice of points, over which the null hypothesis can be rejected (Holmes et al., 1996). Figure 2 shows those areas in the brain with t-values equal or higher than the top 5% of t-values present in the null distribution.

Time-Frequency Analysis: Point of Interest (POI):

In the whole brain analysis, oscillatory signals were strongest and most extensive in the 25-35Hz frequency band, within the 200-400ms time window (see Figure 2 and Supplementary Figure 1) and therefore peaks in these maps were used to identify POIs. Following the selection of POIs from the whole-brain beamforming analysis, separate beamformers (Huang et al., 2004) were used to reconstruct the neural activity in left ATL (MNI coordinates -48,8,-18, in anterior STG), pMTG (MNI coordinates -50,-52,8) and posterior IFG (MNI coordinates -36,16,22). These sites corresponded to points showing the strongest task-induced changes in oscillatory power within the relevant regions of cortex. Although bilateral ATL is implicated in semantic representation, we focused on left-hemisphere sites since (i) the stimuli were written words; (ii) fMRI and patient studies reveal a greater contribution of the left hemisphere to semantic processing, especially for written words (Binder et al., 2009; Rice, Lambon Ralph & Hoffman, 2015; Noonan et al., 2013); and (iii) right motor cortex was expected to show irrelevant responses related to the preparation of button presses with the left hand (even though button presses were only required on catch trials), and therefore contaminate the signals of interest.

We then used the Stockwell transform (Stockwell, Mansinha, & Lowe, 1996) to calculate time-frequency representations for each POI from 5-50 Hz over the time period -800 to 700 ms, where 0ms was the onset of the second word. This allowed us to examine the response to semantic matching from 0-600ms, with reference to a passive period before the onset of the first word (defined as -700 to -500 ms as in the whole brain analysis). The Stockwell transform, implemented in the NAF software, uses a variable window length for the analysis which is automatically adapted along the frequency range according to the sample rate and the trial length (4th order Butterworth filters with automatic padding). The time-frequency representations of total power were normalized, separately for each condition and for each participant, with respect to the mean power per frequency bin in a baseline period prior to the start of trials in that condition (-700 to -500 ms). This window length was also used in earlier studies (Mollo et al., 2017; Wheat et al., 2010; Klein et al., 2014), since it provides a compromise between the minimum length sufficient to estimate power at the lowest frequency reported here (i.e., 5Hz) and the requirement to characterise the state of the brain immediately before the onset of each trial.

To compare the time-frequency representations between experimental conditions, we used PROC MIXED in SAS (SAS Institute Inc., North Carolina, US) to compute generalized linear mixed models (GLMM). Time-frequency plots of percentage signal change were treated as two dimensional arrays of small time-frequency tiles, indexed in the model by three main effects: time, frequency and the interaction between time and frequency. Therefore, random effects were included in each GLMM to account for the fact that each participant's time-frequency plot is made up of multiple time-frequency tiles. Time-frequency (or spatial) co-variance in the spectrogram was controlled for by assuming the estimates of power followed a Gaussian distribution: consequently a Gaussian link function was used in the model. The time-frequency (spatial) variability was integrated in the model by specifying an exponential spatial correlation model for the model residuals (Littel et al., 2006). Finally, the data were resampled at a frequency resolution of 2.5Hz and time resolution of 25ms, the smallest time and frequency bin consistent with model convergence. This time-frequency resolution proved optimal in other similar published studies (Mollo et al., 2017; Klein et al., 2014; Urooj et al., 2014; Wheat et al., 2010). PROC MIXED constructs an approximate *t* test to examine the null hypothesis that the LS-Mean for percentage signal change between conditions was equal zero in each time-frequency tile, and the procedure automatically controls for multiple comparisons (i.e. controlling experiment-wise type I error). The statistical contours on the percentage signal change figures for total power encompass time-frequency tiles fulfilling both of the following criteria: a) the difference between conditions reached $p < 0.05$; b) any region in the time-frequency plot defined by (a) also showed a response that was significantly different from zero in at least one of the two contributing conditions.

Results

Behavioural experiment

While traditional priming experiments show facilitation for weakly-related as well as strongly-related primes, compared with unrelated words (Neely, 1977; 1991), weak associations are expected to show a processing cost when making explicit semantic decisions (Badre et al., 2005; Whitney et al., 2011). The behavioural data were consistent with these predictions (Figure 1b and 1c). A one-way repeated-measures ANOVA of reaction times from the behavioural pre-scan results showed a statistically significant main effect of experimental condition ($F(2,38) = 22.26$, $p<.001$; Figure 1b). Post-hoc comparisons showed that reaction times were faster for strong associations compared with both the weak and unrelated conditions ($t(38)=6.25$, $p<.001$ and $t(38)=5.15$, $p<.001$ respectively). There was no statistically significant difference in reaction times between the weak and unrelated conditions. A similar analysis for accuracy showed a statistically significant main effect of condition ($F(2,38) = 31.47$, $p<.001$; Figure 1c). Post-hoc comparisons showed that accuracy for the weak condition was significantly lower than that for both the strong and unrelated conditions ($t(38)=6.78$, $p<.001$ and $t(38)=6.95$, $p<.001$). There was no significant difference in accuracy between the strong and unrelated conditions within pre-scan behavioural experiment.

Reaction times were generally longer for catch-trials recorded during MEG acquisition, perhaps because participants did not always reach an explicit decision until the requirement to respond was indicated. Nevertheless, the data followed a similar pattern to the pre-scan experiment. A one-way repeated-measures ANOVA of reaction times from the catch-trials in MEG showed a statistically significant main effect of experimental condition ($F(2,38)=10.63$, $p<.001$), as shown in Figure 1b. Post-hoc comparisons showed faster reaction times for strong associations compared with both the weak and unrelated conditions ($t(38)=4.60$, $p<.001$ and $t(38)=2.50$, $p<.05$ respectively). In addition, reaction times for the unrelated condition were significantly faster than those for the weak condition ($t(38)=2.10$, $p<.05$). A similar analysis of catch-trial accuracy showed a main effect of condition ($F(2,38)=89.03$, $p<.001$), as shown in Figure 1c. Post-hoc comparisons showed that accuracy for the weak condition was significantly lower than that for both the strong and unrelated conditions ($t(38)=11.66$, $p<.001$ and $t(38)=11.45$, $p<.001$). However, there was no significant difference in accuracy between the strong and unrelated conditions.

Whole-brain results

The response to the task as a whole (i.e., the response to the second word of the pair collapsed across both experimental conditions, versus a period prior to the start of the trial), is shown in Figure 2. The most extensive changes in total power in response to the task were power decreases, relative to the resting passive period, in the 25-35Hz frequency band (shown in Figure 2 below). Other frequency bands showed similar effects of the task in the temporal lobe but only the 25-35Hz frequency band showed a response in anterior cortical regions (see supplementary Figure 1). These decreases in total oscillatory power were focussed on temporal, occipital, inferior frontal and parietal lobe regions implicated in visual and semantic processing, starting within the first 200ms and lasting for at least 600 ms after target presentation. Decreases in total power are commonly interpreted as reflecting an *increase* in neural activity that is not phase-locked to stimulus presentation (Hanslmayr et al., 2012). Reductions in total power have been shown to correlate with an increased BOLD response in fMRI (Hanslmayr et al., 2011; Singh et al., 2002; Hall et al., 2014), and a recent review proposed that decreases in total power reflect active engagement of neocortex in the encoding and retrieval of memories (Hanslmayr, Staresina & Bowman, 2016). Thus, the whole-brain beamforming results are consistent with an increase in visual and semantic processing following the onset of the second word.

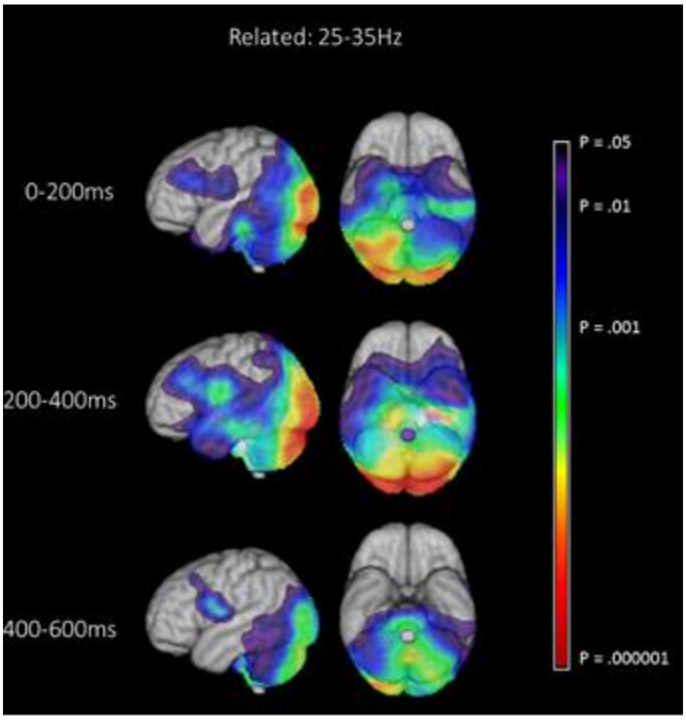


Figure 2: Whole-brain beamforming results for the 25-35 Hz frequency band, showing differences in total oscillatory power between an active period following target onset and a passive period prior to each trial. The first 600ms following presentation of target word are displayed, in 200ms windows. Task effects were decreases in total power in all cases. The images show a t-value map, thresholded at p<.05.

Points of interest results

Whole epoch data for each site

For each POI, Figure 3 shows time-frequency plots of total power for the whole epoch, corresponding to the first and second word responses in each semantically-related pair. These plots are included to illustrate the response to the task at each site, and to inform the interpretation of contrasts between conditions that were computed from the onset of the second word, in the context of ongoing task activity. Orange-red-brown colours indicate *power increases*, whereas green-purple-black colours indicate *power decreases* relative to the baseline (with no change shown in green). In all three sites, there was a subtle increase in oscillatory power in response to the first word, while the presentation of the second word was characterised by a large *reduction* in total oscillatory power relative to baseline. The reduction in oscillatory power followed the offset of the first word in pMTG, anticipated the onset of the second word in IFG, and followed the onset of the second word, building over time in ATL.

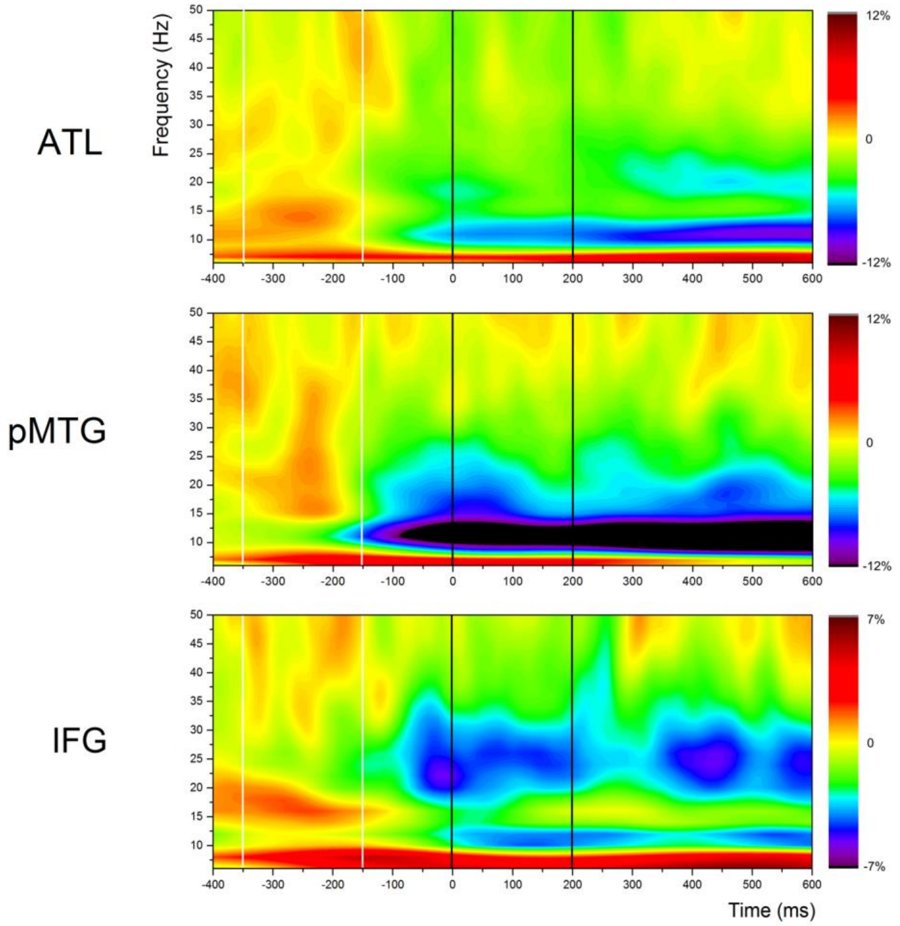


Figure 3: Total oscillatory power across the whole epoch for related trials, including both words presented in the relatedness judgement task. Presentation of the prime word (first word of the pair) is shown within white vertical lines, while presentation of the target word (second word of the pair) is illustrated within black vertical lines. Orange-brown indicates regions of *power increase* relative to the baseline, while green-purple indicates *power decreases* relative to the baseline, and yellow indicates no change from baseline

Differences between conditions in POIs

As shown in Figure 4c, we found statistically significant differences between strong and weak associations throughout the epoch, in the beta and low gamma frequency bands, in all three sites. However, strength of association had opposite effects at the two temporal lobe sites, in line with the predictions of the Controlled Semantic Cognition framework (Lambon Ralph et al., 2017). ATL showed a greater change from baseline during the retrieval of strong vs. weak associations. This effect was significant from 400ms post-target onset until the end of the epoch at 7-12 Hz. PMTG, in contrast, showed stronger changes in oscillatory power during the retrieval of weak associations, from within 100ms of the onset of the second word, and this effect lasted throughout the epoch (to 550ms, at around 15Hz, plus brief significant differences at 25Hz and 30Hz). Like pMTG, IFG also showed stronger changes in oscillatory power during the retrieval of weak than strong associations soon after the onset of the second word: there were greater power increases for this condition at a very low frequencies (below 10Hz) from 0-200ms, and stronger task-induced decreases in power at 25Hz and 50ms post-onset of the second word. However, from around 200ms, this response reversed, such that task-induced decreases in total power were greater for strong associations in IFG from 15-20Hz.

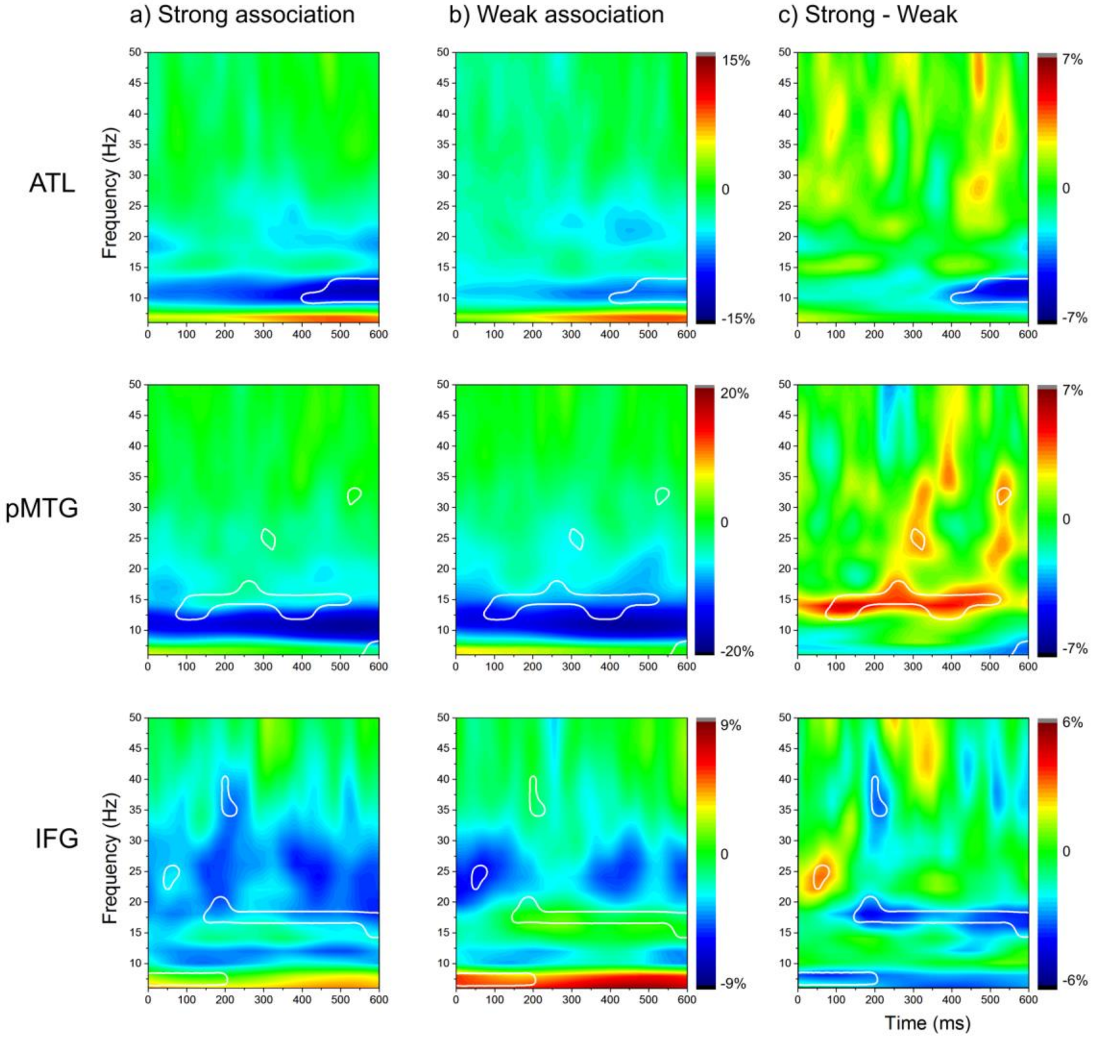


Figure 4: a) Percentage signal change in the strong condition, relative to baseline. b) Percentage signal change in the weak condition, relative to baseline. c) Percentage signal change between strong and weak conditions, separately for ATL, pMTG and IFG. White lines are derived from the statistical

comparison between strong and weak conditions. The boundaries enclose regions fulfilling two criteria: i) percentage signal change between the strong and weak conditions is significantly different from zero (p<0.05) and ii) percentage signal change computed separately for each condition is significantly different from zero for at least one of the two conditions. Yellow-red colours indicate regions of *power increase* relative to the baseline, while cyan-blue indicates *power decreases* relative to the baseline, and green indicates no change from baseline.

While the focus of this study was on differences between strong and weak associations to test the predictions of the Controlled Semantic Cognition Framework (Lambon Ralph et al., 2017), we also computed differences between related and unrelated trials to allow comparison with previous studies that employed similar contrasts (for a review, see Lau et al., 2008). The results of this analysis can be seen in Supplementary Materials.

Summary of MEG results

Comparisons of strong and weak associations revealed a dissociation in the temporal lobe, in both space and time, which depended on the match between the semantic retrieval required by the task and the structure of long-term conceptual knowledge. ATL showed a strong response during the retrieval of both strong and weak associations soon after the presentation of the second word, plus greater oscillatory power for strong than weak associations from around 400ms after target onset. This is consistent with the view that ATL supports coherent semantic retrieval when inputs and task requirements align with long-term conceptual representations (Feng et al., 2016; Binder, 2016). The timing of this result is consistent with previous studies showing strong semantic effects in ATL around 400ms and suggests that effects of coherent semantic retrieval emerge over time (Jackson, Hoffman, Pobric, & Lambon Ralph, 2016; Lau et al., 2013; 2014; Marincovic et al., 2003). In contrast, pMTG and IFG both showed greater oscillatory power for weak than strong associations soon after the onset of the second word, suggesting a role in detecting circumstances where inputs are inconsistent with the semantic context elicited by the first word (triggering the recruitment of controlled retrieval processes). Given the rapid response in pMTG to the offset of the first word and the sustained engagement at this site for weak > strong associations throughout the trial, this site might also play a role in maintaining aspects of knowledge that are currently relevant.

Experiment 2: Chronometric TMS

Experiment 1 demonstrated a dissociation in oscillatory power within the temporal lobe in space and time, reflecting the extent to which the pattern of semantic retrieval required by the task was consistent with dominant aspects of long-term conceptual knowledge. To determine the causal role of ATL and pMTG in semantic retrieval, Experiment 2 used chronometric TMS to disrupt processing in these two regions at different points in time in the same paradigm. Stimulation of ATL during the presentation of the second word in the pair might disrupt the efficient retrieval of strong associations, given the MEG findings above. In addition, stimulation of pMTG at an earlier time-point after the onset of the second word might be expected to disrupt the retrieval of weak more than strong associations.

Materials and Methods

Participants:

Participants were 15 right-handed native English speakers, with normal or corrected-to-normal vision, and no history of language disorders (8 males, mean age 23, age range 20-32 years). This experiment employed a separate sample from Experiment 1. Written consent was obtained from all participants and the study was approved by the York Neuroimaging Centre Research Ethics Committee.

Design

The experiment employed a 3x2x4 repeated-measures design, with site (ATL, pMTG and sham mid-MTG), task (semantic association task and digit parity judgement task), and TMS timings (0 & 40ms; 125 & 165ms; 250 & 290ms; 450 & 490ms) as within-subject factors. At each time point, a pair of pulses 40ms apart was applied, since this dual-pulse method is thought to generate more significant behavioural disruption than single pulses (Gagnon, Schneider, Grondin & Blanchet, 2011; Strafella & Paus, 2001; Chen, 2000). The stimulation times were selected to provide coverage of time points of interest from the MEG experiment: these included processes already in play by the onset of the second word, which are likely to be important given the successive stimulus presentation used in our paradigm, responses observed 100-200ms after the onset of the second word (by which point the differential response in pMTG was established), effects within the first 300ms (e.g., related > unrelated differences in ATL), and later effects. This allowed us to explore these sites' causal involvement in retrieving dominant and weaker aspects of knowledge.

Materials

The semantic task was the same as for Experiment 1. Word pairs were presented sequentially, and participants decided whether the two words were related or not. The pairs were either strongly or weakly associated, or they were unrelated. To maximise sensitivity to the effects of TMS on the retrieval of strong and weak associations, each session comprised 70% related trials (which were the focus of the analysis) and 30% unrelated trials (to ensure participants attended to the task, which were excluded from the analysis). The same target words were presented across conditions, although each target was only presented once per session. In addition, the first words of the strong and weak pairs did not differ in word frequency or length (see Table 2).

Table 2: Comparing word frequency and length for the first word across conditions, plus the associative strength between the two words in the TMS experiment

Measure	Strong Association M (SD)	Weak Association M (SD)	p-value
Word frequency	17.43 (32.38)	19.28 (32.91)	.66
Word length (letters)	5.62 (1.81)	5.48 (1.51)	.49
Association strength	0.43 (0.19)	0.03 (0.06)	.001

A non-semantic task involving numerical judgements was designed to match the semantic task in overall difficulty. Previous fMRI and TMS experiments have employed similar numerical control tasks (e.g., Pobric et al., 2007) because number representations are thought to be independent of temporal lobe semantic regions. We therefore used this task in an attempt to control for non-specific effects of TMS. Two three-digit numbers were presented sequentially, and subjects were asked to decide whether *both numbers* were odd or even. The proportion of yes/no trials was identical to the semantic task (i.e., 70% match). One participant was tested on a different number judgement task and was excluded from the statistical comparisons of semantic vs. number task performance. For the word conditions, there were 25 trials with TMS delivered to each of the three stimulation sites at 4 different timings (25×4×3), for each condition (strongly related, weakly related, unrelated). For the digit task, there were 12 trials for each of the three stimulation sites at 4 different timings (12×4×3), for each number “condition” (both even, both odd, different).

Stimulus presentation

The three experimental sessions were divided into 5 runs, each lasting approximately 12 minutes. TMS was delivered in 4 of the 5 runs, and a block without TMS was placed in the middle of the 5 runs for safety reasons. Each run was made up of 6 blocks for each task (numerical or semantic), lasting around 60 seconds. Blocks were arranged in pseudorandom order to minimise task switch costs. When switching between tasks, a short instruction screen informed the participant which task would be presented next. The first trial after the task switch was a dummy trial which was discarded from further analysis. The first word of the pair was presented for 200ms, followed by an inter-stimulus interval (ISI) of 150ms, and then the second word requiring a relatedness judgement appeared for 500ms (see Figure 1a). The nonius lines remained on screen for 1000ms, and were then dimmed for 1150ms after the participant's response, to signal the end of the trial. Following this, the bright nonius lines returned, to cue the onset of the next trial, for a randomly variable interval of 0-1000ms (500ms on average) before the onset of the first word of the next pair. Each trial lasted on average 3500ms. As in the MEG experiment, participants were asked to decide if the two words were related in meaning or not. They responded with their right hand and were instructed to be as quick and accurate as possible. Before starting the experiment, participants performed a practice session with 10 trials of both tasks (without TMS), and three practice trials with stimulation. Participants took self-paced breaks between the runs.

Stimulation sites

TMS was applied to left ATL, left pMTG, and a sham site in the mid-temporal lobe (halfway between these two sites). Stimulation sites were taken from published studies; participants' structural T1 MRI scans were co-registered to the scalp using the Brainsight frameless stereotaxy system (Rogue Research, Montreal, Canada) to identify the stimulation targets in each participant's brain. The left ATL site was in anterior ventrolateral temporal cortex (MNI -51, 6, -39; coordinates from Binney et al., 2010). This site showed greater activation for synonym judgement than numerical magnitude judgement in fMRI, and is located close to the region of peak atrophy in semantic dementia (Binney et al., 2010). The left ATL coordinates for TMS fell within the area of statistically significant oscillatory power revealed by whole-brain beamforming in Experiment 1, although the peak in the MEG data was anatomically superior (21 mm), and somewhat lateral and anterior (3 mm and 2 mm respectively) relative to the stimulation site. Similarly, the choice of left pMTG site for TMS was based on a meta-analysis of neuroimaging studies of semantic control by Noonan et al. (2013; MNI -58, -50, -6). This site activates across a wide range of manipulations of semantic control, and shows a stronger response to weak than strong associations (Davey et al., 2016; Gold et al., 2006). It was also located within the area of statistically significant oscillatory power revealed by whole-brain beamforming in Experiment 1, but was inferior (14 mm) and lateral (8 mm) to the pMTG POI. We opted to use stimulation sites from the literature rather than peaks from Experiment 1 given the relatively poor spatial resolution of MEG. The sham control site was selected by finding the midpoint on the y-axis between the two experimental sites, varying the z coordinate to ensure that stimulation was delivered to the middle temporal gyrus.

TMS stimulation protocol

Chronometric TMS was delivered using a Magstim Rapid2 stimulator and a 50mm diameter figure-eight coil. Stimulation intensity for ATL and pMTG was 60% of the maximum output of the stimulator. We did not measure the intensity of stimulation required to elicit a visible hand twitch (i.e., active motor threshold) as it is unclear whether this predicts excitability at stimulation sites far from motor cortex (Antal et al., 2004; Gerwig et al., 2003). The sham stimulation was applied at 30% of stimulator output since this intensity is thought to be too weak to produce a neural effect, but it still mimics the sound and scalp sensations of TMS stimulation (Duecker et al., 2013). Dual-pulse TMS was delivered at 25Hz (pulses 40 ms apart) in each trial (see Figure 1a for illustration). The position of the coil was monitored and tracked in real time. The mean difference between the intended target and the stimulated site on each trial was 0.3mm (s.d. = 0.26; maximum displacement = 5.6mm). Trials in the different timing conditions were arranged in an ascending or descending staircase of 4 trials (i.e. four trials with stimulation at 0 & 40ms followed by four trials of stimulation at 125 & 165ms etc.). We used this strategy to limit the participants' awareness of the different TMS timings, and to reduce any tendency to wait until stimulation had been delivered before responding (Sliwinska et al., 2012). Following safety guidelines (Rossi et al., 2009), an inter-train interval of 5000ms was added after every sequence of 24 double pulses. Where possible this interval corresponded to the task switching instruction screen; in other cases it was added after a button press response.

Analysis strategy

We wanted to know how speeded judgements about strong and weak semantic relationships between pairs of words would be affected by TMS, delivered at different time points following the onset of the second word in a pair, at the two different cortical sites. To maximize the sensitivity of these analyses, we used generalised linear mixed models (GLMM) which retained information about all trials and permitted random effects at both the participant and item levels to be modelled (see Baayen, Davidson & Bates, 2008). To do this, we specified an 'unstructured' variance-covariance structure for each random effect in the model's G-matrix. The mixed models were implemented in PROC MIXED in SAS v9.4 (SAS Institute, North Carolina, USA).

Previous TMS studies have reported consistent slowing for semantic decisions following inhibitory stimulation, and little effect on accuracy (Walsh & Cowey, 2000; Pasqual-Leone, Walsh & Rothwell, 2000; Devlin, Matthews & Rushworth, 2003). Therefore, our primary outcome variable for each trial was the magnitude of the TMS effect, defined as the difference in response time between a word pair subject to TMS and its corresponding sham version. Incorrect responses and outlying data points that fell more than 2SD from each participant's mean RT were removed, for each session, prior to analysis.

For the initial models, we included the main effects of task condition (e.g., strong vs. weak association), site (ATL, pMTG), and TMS time (i.e., pulses at 0-40ms; 125-165ms; 250-290ms; 450-490ms after the onset of the second word), plus their interactions. We also included as covariates structural aspects of the experiment (i.e. session and block order). In addition, supplementary analyses, characterising (i) the effect of TMS on accuracy for strong and weakly-related targets and (ii) the effect of TMS on semantic judgements overall (vs. numerical judgements), highlighted non-specific effects of TMS on both RT and accuracy in our data, as we report in the Supplementary Materials. For these reasons, accuracy per block and performance in the numerical task were also included as covariates in the initial models for reaction time. The criteria we used to optimize the final model were: (i) a significant reduction in -2Log-Likelihood relative to the empty model, (ii) only explanatory variables that were statistically significant at p<.05 should be retained. Once the final model was fitted, we used PROC MIXED to estimate pairwise t-test comparisons of the least squared (LS) mean reaction times, with and without TMS, carried out separately at each site for each condition (a total of 5041 observations). These post-hoc comparisons were controlled for multiple comparisons.

Results

The main effects from the optimized GLMM of reaction time are shown in Table 3. Since our dependent measure was the TMS effect (computed as the difference between TMS and sham trials), there was no main effect of condition. We found significant main effects of TMS time (reflecting greater differences between stimulation and sham at later time points) and site (reflecting a greater difference between ATL and sham than between pMTG and sham). These effects are likely to be explained by non-specific effects of stimulation. Larger TMS effects for ATL than for pMTG might reflect the distracting effects of the strong temporalis muscle contractions that occur during stimulation of this site. Moreover, a similar main effect of time has been observed in previous chronometric TMS studies (Sliwinska et al., 2012) and is thought to reflect a tendency for participants to wait before responding on trials in which the TMS pulse is applied comparatively late (see Supplementary Analysis 3 for further discussion). The covariates of block, session order, and number RT (characterising non-specific effects of TMS) were also statistically significant, although the accuracy covariate did not improve model fit and was not included in the final model. Critically, there was a significant three-way interaction between condition (strong vs. weak), TMS time and site, suggesting that the disruption of strong and weak associations occurred at a different point in time after the onset of the second word of the pair, and that this effect was different comparing ATL with pMTG.

Table 3: Effect of TMS on RT for strong and weak associations

Model Parameter	F-value (d.f.)	Z-value	p-value	-2Log likelihood
Empty model				63784.3
Time	4.86 (3, 302)		.0026	63277.5
Site	12.39 (1, 4783)		<.001	
Condition × Time × Site	1.95 (11, 4920)		.029	
Block order	15.16 (3,5004)		<.001	
Testing session	4.41 (2, 4465)		.012	
Number task RT	191.9 (1,1955)		<.001	
Participant covariance		2.59	.0048	
Target covariance		4.75	<.001	

Figure 5 shows mean reaction times (upper row) and the post-hoc comparisons of LSmean reaction times (low row), separately for ATL and pMTG. For ATL, we found a significantly larger effect of TMS on strongly-related than weakly-related pairs (giving rise to a positive LSmean difference in the bottom row of Figure 5), when pulses were applied at 125-165ms after the onset of the second word. At the other time points, the magnitude of the TMS effect was equivalent for the strong and weak associations. This suggests that at around 150ms post-presentation of the second word, the efficient retrieval of strong semantic relationships was disrupted by the perturbation of ongoing processing within ATL. Although strong associations did not evoke a stronger change in oscillatory response at this site until later (400ms in the MEG data), and the behavioural response was later still (between 500-600ms in this experiment), disruption of a settling process within ATL might potentially disrupt or delay both of these subsequent effects.

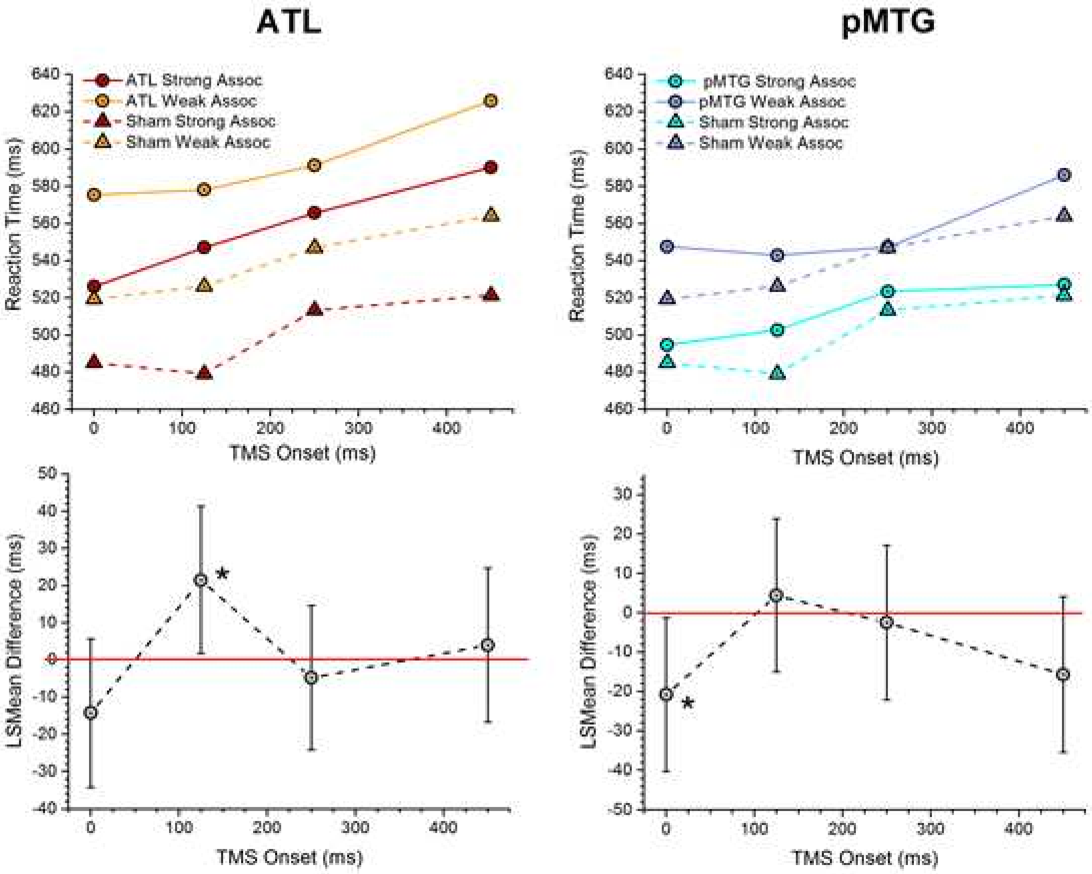


Figure 5: Effect of TMS on RT for strong and weak associations. TOP ROW: RT (in ms) for the strong and weak conditions for ATL (left) and pMTG (right). RT data for the strong and weak condition for the sham site is showed in dashed lines. These plots show the raw (un-modelled) means. BOTTOM ROW: A comparison of LS Means differences between strong and weak conditions in the effect of TMS. Data points above the red line indicate greater disruption for the strong condition, while data points below the red line indicate greater disruption for the weak condition. Statistically significant differences (at p<.05) between the effects of TMS on strong and weak trials are indicated with asterisks. Error bars show 95% confidence intervals.

For pMTG, we found a significantly larger effect of TMS on weakly-related than strongly-related pairs (giving rise to a negative LSmean difference in the bottom row of Figure 5), when pulses were applied at 0-40ms after the onset of the second word in the pair. At the other time points, the magnitude of the TMS effect was equivalent for the strong and weak associations. This very early differential response suggests that pMTG may make a critical contribution to the capacity to *engage* controlled retrieval when it is needed. Stimulation at this early point may have disrupted the maintenance of current contextual information generated by the first word in the pair. This could disproportionately affect weak associations if, for example, pMTG plays a critical role in detecting the need to employ controlled retrieval. Although weak associations did not evoke a stronger change in oscillatory response at pMTG until slightly later (from around 60ms in the MEG data), effects linked to controlled retrieval at pMTG in both MEG and TMS were observed very early after the onset of the second word, allowing us to reject one view of the emergence of semantic retrieval over time, in which conceptual knowledge is first activated or retrieved and then subsequently selected to suit current task demands or the context.

Discussion

A significant body of research has characterised the brain regions that support semantic processing but less is known about the temporal evolution of semantic retrieval across these regions. While studies have examined the time course of semantic access from written words and pictures following a semantically-related or an unrelated item (Dikker & Pylkkänen, 2013; Halgren et al., 2002; Lau et al., 2013; 2014), the focus here was on the brain processes that support the explicit retrieval of strong associations (which are expected to be supported by their coherence with the structure of long-term semantic knowledge) as opposed to weak associations (which are less well-supported by long-term conceptual information and thus might require greater engagement of controlled retrieval processes to shape retrieval to suit the demands of the task). We examined how the retrieval of strong and weak semantic conceptual relationships was reflected in (i) changes in oscillatory power over time, as measured by MEG; and (ii) vulnerability to inhibitory online brain stimulation, using chronometric TMS.

In both experiments, the same behavioural paradigm was used to explore the functional and temporal organisation of semantic processing in the left anterior and posterior temporal lobe (ATL and pMTG) and inferior frontal gyrus (IFG; in the MEG experiment only; this site was not stimulated with cTMS). Previous work has associated ATL with the retrieval of strong associations, in conjunction with other regions in the default mode network (Davey et al., 2016; Jackson et al., 2015), while controlled retrieval is thought to engage semantic control processes in left pMTG, together with IFG, to allow non-dominant aspects of meaning to come to the fore (Noonan et al., 2013; Badre et al., 2005; Gold et al., 2006; Davey et al., 2015; 2016). In line with these predictions, task-induced changes in oscillatory power were greater for strong than weak associations in ATL, while pMTG showed the opposite pattern (weak > strong associations). TMS confirmed a causal role for these sites in the efficient retrieval of strong and weak associations respectively. IFG initially showed stronger oscillatory power for weak associations, suggesting that this site contributes to the establishment of a suitable network for semantic retrieval (Mollo et al., 2018, in press) but later this effect was reversed, suggestive of a more general role in semantic retrieval.

Timing differences between the sites were also found: ATL showed greater oscillatory power for the strong associations around 400ms post-target onset, although a strong task-related response was observed in the MEG data across conditions even before the onset of the second word (reflecting the successive presentation of multiple meaningful items in our paradigm). TMS to ATL disrupted performance for strong associations at around 150ms, around the time that early effects of semantic manipulations have been reported at this site in other studies (Clarke et al., 2011; 2012; Hauk et al., 2006). This time point may have been sensitive to the disruptive effects of TMS (even though the difference between strong and weak conditions was not significant in the MEG data until later) since a coherent pattern of semantic retrieval was not yet fully established (and was therefore vulnerable to interference). pMTG showed an even earlier differential response to the strong and weak conditions in both MEG and TMS: this site responded more strongly to weak associations throughout the analysis window (from about 60ms post-onset of the second word), and TMS delivered to pMTG at the point of target onset impaired the efficient retrieval of weak associations. Thus, the MEG and TMS results followed the same temporal sequence across sites, although the critical time for TMS-induced disruption preceded the emergence of condition differences in MEG. IFG also showed an early response to weak associations, although this effect was not sustained as it was for pMTG, and we did not investigate critical time-points for IFG using cTMS. Below, the contributions of left ATL, pMTG and IFG to semantic cognition are discussed in light of these findings.

Anterior temporal lobe: The ATL is proposed to play a crucial role in heteromodal conceptual representation (alongside modality-specific 'spokes'; Patterson, Nestor & Rogers, 2007; Rogers et al., 2006; Coutanche & Thompson-Schill, 2014). ATL is important for accessing conceptual knowledge from visual inputs (alongside other modalities) – a process that activates the ventral visual stream which terminates in ATL (Visser, Jefferies, Embleton & Lambon Ralph, 2012; Visser, Jefferies & Lambon Ralph, 2009). MEG studies of this aspect of ATL processing have identified responses in this region within 120ms of stimulus onset (Clarke et al., 2013; Fujimaki et al., 2009; Yvert et al., 2012). In addition, ATL is implicated in relatively automatic aspects of semantic access and retrieval (Lau et al., 2013; Davey et al., 2016). The current findings are highly consistent with this emerging story about the contribution of the ATL to semantic processing but add several important elements.

First, we used beamforming to characterise the response in ATL to strong and weak associations in total oscillatory power. In contrast, other MEG studies localising semantic effects to ATL have largely used measures maximally-sensitive to evoked power (Halgren et al., 2002; Bemis & Pylkkänen, 2011; Westerlund & Pylkkänen, 2014; Zhang & Pylkkänen, 2015; Lau et al., 2014; Fujimaki et al., 2009). Total power includes both phase-locked components and signals that are *not* phase-locked to the onset of the stimulus. Since the emergence of coherent semantic activation over time draws on long-term knowledge of the meanings of words across contexts, one might expect this process to generate neural oscillations that are not directly linked to stimulus onset. In line with these considerations, strong task-induced decreases in total power to the second item were found in all three sites. These effects were not seen in response to the presentation of the first word in the pair (see Figure 3), and therefore this response could be a marker of meaning retrieval that is at least partly decoupled from the stimulus itself. This interpretation draws on the view that power decreases are not necessarily associated with a decrease in neural activity (Hanslmayr et al., 2012; Hanslmayr, Staresina & Bowman, 2016): decreases in total power can reflect an increase in desynchronised neural activity that allows the representation of richer informational content, and our results can be interpreted within this framework – strong associations are more coherent with the structure of long-term conceptual knowledge and might generate richer or more meaningful experiences in ATL.

TMS to ATL disrupted the efficient retrieval of strong more than weak associations at 150ms post-stimulus onset – i.e., at the point when interactions between visual cortex and ATL are thought to become established (Clarke et al., 2011; 2012). In the MEG data, there was a strong task-related response in ATL by 150ms, although there was not yet a significant difference between the strong and weak conditions. Thus, the emergence of coherent semantic retrieval for the strongly-linked items may have been vulnerable to perturbation from TMS before the pattern of response within the ATL was well-established. Although a previous cTMS study found disruption when TMS pulses were applied to ATL at 400ms post-trial onset (Jackson et al., 2015), this study did not examine differential disruption of strong vs. weak associations, and it involved a more complex two-alternative-forced-choice decision as opposed to yes/no decisions about the presence or absence of a relationship between two words – thus the timings are unlikely to be comparable.

Posterior middle temporal gyrus and inferior frontal gyrus: The involvement of left IFG in semantic control is relatively well-established (e.g., Badre et al., 2005; Noonan et al., 2013), yet there is considerable controversy about the role of left pMTG in semantic cognition, since dominant theoretical frameworks have suggested that this site (i) represents particular aspects of lexical or semantic knowledge – such as event representations; or (ii) supports controlled semantic cognition as part of a large-scale network that includes IFG (see Davey et al., 2016). Studies have shown a common response in left pMTG and IFG using a wide range of manipulations of semantic control – including contrasts of ambiguous over non-ambiguous words, decisions with strong vs. weak distracters, and the retrieval of weaker versus stronger semantic links, in paradigms similar to the one adopted here (Noonan et al., 2013). pMTG is functionally connected to both the executive network and ATL, suggesting this region may be well-placed to control retrieval from the semantic store (Davey et al., 2016). Offline TMS studies have provided convergent evidence for the disruption of weak (but not strong) semantic association judgements when inhibitory stimulation is applied to pMTG as well as IFG (Whitney et al., 2011; Davey et al., 2015). When the relationship between the two words is weak, the first word in the sequence will tend to activate features and associations that are irrelevant to the decision that has to be made, and consequently we expect controlled retrieval processes will be triggered to shape the emerging pattern of semantic retrieval so that it focusses on the relevant link.

The time-course of these effects place important constraints on theories of controlled semantic retrieval: IFG and pMTG would be expected to show a relatively late response to the comparison of weak and strong if controlled retrieval takes time to become established, and if activity at this site reflects a re-interpretation or re-shaping of semantic activation following initial semantic retrieval driven by the written input. Alternatively, these semantic control regions might show an early response to the same comparison if they are important for triggering the recruitment of the semantic control network when incoming information is not strongly coherent with ongoing semantic retrieval. This hypothesis is predicated on recent accounts of visual word recognition which suggest extremely rapid interactions between visual, semantic and articulatory codes, as opposed to an orderly sequence of steps from orthography to meaning (Klein et al., 2012; Sereno et al., 2003; Wheat et al., 2010; Woodhead et al., 2014; Yvert et al., 2012). By this view, pMTG and IFG may reduce the propagation of dominant features and associations recovered from ATL when initial processing of new inputs suggests that these aspects of knowledge may be insufficient for comprehension. In addition, these regions might show a sustained response to weak associations if they maintain currently-relevant semantic information which can be used to appropriately constrain activation within ATL.

The current MEG data showed early engagement of both IFG and pMTG that was stronger for weak associations, supporting this alternative interactive view. The weak > strong effect commenced within 50ms of target-onset at both sites implicated in semantic control. This pattern then continued throughout the analysis window for pMTG; however, in IFG, the effect reversed to reveal a strong > weak association pattern by 400ms. Consequently, while both sites might play a crucial role in setting up an appropriate network for controlled retrieval when input processing for the second item was not well-supported by ongoing semantic retrieval to the first item in the pair, pMTG might play a greater role than IFG in *sustaining* a non-dominant pattern of semantic retrieval. cTMS also provided evidence for an early role of pMTG in the efficient retrieval of weak associations, since there was greater disruption of weak trials when TMS was applied at the onset of the second word (when semantic retrieval was also underway but not tuned to the relevant semantic link). These findings are consistent with the hypothesis that pMTG (potentially together with LIFG) maintains currently-relevant features or interpretations and detects situations in which incoming information is not well-aligned with these aspects of knowledge. This interpretation is consistent with studies that have shown a stronger response to more predictive primes in pMTG, including adjectives (Fruchter et al., 2015) and pictures (Dikker & Pylkkänen, 2013) that are informative about upcoming items. In our task, information about the semantic context might have been more critical for the efficient retrieval of weak associations, since it might have supported the rapid engagement of controlled retrieval processes when expectations were partially met. In contrast, for strong associations, relevant features in the semantic store will have been primed by the first word and thus this process may be less critical. If this interpretation is correct, application of TMS even before the onset of the second item may have had a similar effect, since it would have disrupted maintenance of a conceptual 'prediction' that allowed the detection of a situation in which semantic control processes needed to be deployed. However, this remains an untested prediction. This perspective is further consistent with studies suggesting that pMTG shows strong engagement when meaningful inputs themselves determine a context that requires semantic retrieval to be shaped in a particular way (Davey et al., 2016; Badre et al., 2005).

Some limitations of this research are worth noting. First, this study focuses on the role of key locations predicted to show a functional dissociation in the Controlled Semantic Cognition framework (left ATL, pMTG and IFG). By combining targeted analysis of MEG data (examining local peaks within these regions) with chronometric TMS delivered to these sites, strong conclusions can be drawn about the nature of the dissociation in the temporal lobe, although the study is uninformative about other regions in the brain. Secondly, there is increasing evidence of functional subdivisions within these sites. In ATL, temporal pole, ventral ATL and aSTG appear to have different functional profiles (Lambon Ralph et al., 2017; Murphy et al., 2017). There are also functional subdivisions within IFG (Badre et al., 2005; Davey et al., 2016; Noonan et al., 2013), while pMTG lies at the intersection of several networks, including default mode, visual and auditory regions (Braga et al., 2013; Yeo et al., 2011) – which might be critical to its contribution to semantic cognition. The limited spatial resolution of MEG, combined with practical limits on the number of TMS sessions, does not permit the separation of these regions. Our MEG analysis was optimised to characterise the oscillatory dynamics of semantic processing for strong and weak associations for regions that responded robustly to the task but the lack of spatial precision inherent in this data does not allow us to draw specific inferences about specific locations.

Thirdly, it may not be appropriate to directly compare timings across the MEG and TMS experiments, since Figure 1b demonstrates that the behavioural responses recorded within the MEG scanner were considerably slower than those obtained in the laboratory. This may have contributed to differences between our experiments; particularly the earlier effects of strength of association seen in the TMS study relative to the MEG study. More generally, this observation supports the view that it may not be possible to precisely specify the timing of neurocognitive responses, since these timings will critically depend on the task or paradigm that they are measured within. For example, the timing of differential responses to strong associations and weak associations might be influenced by experimental factors such as the stimulus-onset asynchrony (SOA), which is known to modulate the extent to which semantic priming draws on automatic or controlled processes (Gold et al., 2006). This study used brief stimuli presentation (200 ms) and a short SOA (150 ms), in order to limit the impact of factors such as stimulus repetition and the proportion of related to unrelated trials (Neely, 1977; 1991). Furthermore, though the priming literature is relevant to our interpretations, our paradigm is not directly comparable to priming experiments, since we required participants to make an explicit judgement of the relationship between the two words, as opposed to examining the facilitatory influence of meaning on reading. An alternative approach, which we adopted here, is to consider the relative timing of behavioural effects *within a paradigm* which can then be localised to different brain regions.

Taken together, these results indicate dissociable roles of ATL and pMTG in semantic retrieval. ATL and pMTG showed opposite effects of strength of association in a semantic judgement task in both the MEG and cTMS experiments, supporting the proposal that these sites make a differential contribution to more automatic and controlled aspects of semantic retrieval.

Acknowledgements

The research was funded by BBSRC (BB/J006963/1). EJ and NS were supported by the European Research Council (SEMBIND – 283530). JS was supported by the European Research Council (WANDERINGMINDS – 646927).

References

Antal, A., Nitsche, M.A., Kincses, T.Z., Lampe, C., Paulus, W. (2004) No correlation between moving phosphene and motor thresholds: A transcranial magnetic stimulation study. Neuroreport, 15, 297-302.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items. *Journal of Memory and Language*, Special Issue: Emerging Data Analysis, 59(4), 390–412. doi:10.1016/j.jml.2007.12.005.

Badre, D., Poldrack, R. A., Paré-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.

Bedny, M., McGill, M. & Thompson-Schill, S. L. (2008). Semantic Adaptation and Competition During Word Comprehension. *Cerebral Cortex* 18(11), 2574–2585

Bemis, D. K., & Pylkkänen, L. (2011). Simple Composition: A Magnetoencephalography Investigation into the Comprehension of Minimal Linguistic Phrases. *The Journal of Neuroscience*, *31*(8), 2801–2814. http://doi.org/10.1523/JNEUROSCI.5003-10.2011

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. http://doi.org/10.1093/cercor/bhs170

Binder, J. R., Desai, R. H. Graves, W. W. & Conant, L. L. (2009). Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex* 19(12), 2767–2796.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M. & Cox, R. W. (1999). Conceptual Processing during the Conscious Resting State: A Functional MRI Study. *Journal of Cognitive Neuroscience* 11(1), 80–93. doi:10.1162/089892999563265.

Binder, J. R., McKiernan, K. A., Parsons, M. E., Westbury, C. F., Possing, E. T., Kaufman, J. N., & Buchanan, L. (2003). Neural Correlates of Lexical Access during Visual Word Recognition. *Journal of Cognitive Neuroscience*, *15*(3), 372–393. https://doi.org/10.1162/089892903321593108

Binney, R. J., Embleton, K. V., Jefferies, E., Parker, G. J. M., & Ralph, M. A. L. (2010). The Ventral and Inferolateral Aspects of the Anterior Temporal Lobe Are Crucial in Semantic Memory: Evidence from a Novel Direct Comparison of Distortion-Corrected fMRI, rTMS, and Semantic Dementia. *Cerebral Cortex*, *20*(11), 2728–2738. http://doi.org/10.1093/cercor/bhq019

Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P. & Hodges, J. R. (2000). Non-Verbal Semantic Impairment in Semantic Dementia. *Neuropsychologia* 38(9): 1207–15. doi:10.1016/S0028-3932(00)00034-8.

Braga, R.M., Sharp, D.J., Leeson, C., Wise, R.J.S., Leech, R. (2013) Echoes of the Brain within Default Mode, Association and Heteromodal Cortices. *Journal of Neuroscience*, 33,14031-14039.

Bright, P., Moss, H., & Tyler, L. K. (2004). Unitary vs multiple semantics: PET studies of word and picture processing. *Brain and Language*, *89*(3), 417–432. http://doi.org/10.1016/j.bandl.2004.01.010

Brookes, M. J., Vrba, J., Robinson, S. E., Stevenson, C. M., Peters, A. M., Barnes, G. R., Hillebrand, A & Morris, P. G. (2008). Optimising experimental design for MEG beamformer imaging. *NeuroImage*, *39*(4), 1788–1802. http://doi.org/10.1016/j.neuroimage.2007.09.050

Brookes, M. J., Wood, J. R., Stevenson, C. M., Zumer, J. M., White, T. P., Liddle, P. F., & Morris, P. G. (2011). Changes in brain network activity during working memory tasks: A magnetoencephalography study. *NeuroImage*, *55*(4), 1804–1815. http://doi.org/10.1016/j.neuroimage.2010.10.074

Brown, C., & Hagoort, P. (1993). The Processing Nature of the N400: Evidence from Masked Priming. *Journal of Cognitive Neuroscience*, *5*(1), 34–44. https://doi.org/10.1162/jocn.1993.5.1.34

Chan, A. M., Baker, J. M., Eskandar, E., Schomer, D., Ulbert, I., Marinkovic, K., Cash, S. S. & Halgren, E. (2011). First-Pass Selectivity for Semantic Categories in Human Anteroventral Temporal Lobe. *Journal of Neuroscience* 31(49), 18119–29. doi:10.1523/JNEUROSCI.3122-11.2011.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*(10), 913–919. https://doi.org/10.1038/13217

Chen, R. (2000). Studies of human motor physiology with transcranial magnetic stimulation. *Muscle & Nerve*, *23*(S9), S26–S32. http://doi.org/10.1002/1097-4598(2000)999:9<::AID-MUS6>3.0.CO;2-I

Chen, Y., A. Shimotake, R. Matsumoto, T. Kunieda, T. Kikuchi, S. Miyamoto, H. Fukuyama, R. Takahashi, A. Ikeda, & M.A. Lambon Ralph. "The 'when' and 'where' of Semantic Coding in the Anterior Temporal Lobe: Temporal Representational Similarity Analysis of Electrocorticogram Data." *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 79 (June 2016): 1–13. doi:10.1016/j.cortex.2016.02.015.

Clarke, A., Taylor, K. I., Devereux, B., Randall, B., & Tyler, L. K. (2012). From Perception to Conception: How Meaningful Objects Are Processed over Time. *Cerebral Cortex*.

Clarke, A., Taylor, K. I. & Tyler, L. K. (2011) 'The Evolution of Meaning: Spatio-temporal Dynamics of Visual Object Recognition'. *Journal of Cognitive Neuroscience* 23(8), 1887–1899.

Corbett, F., Jefferies, E., & Lambon Ralph, M. A. (2009). Exploring multimodal semantic control impairments in semantic aphasia: Evidence from naturalistic object use. *Neuropsychologia*, *47*(13), 2721–2731. Http://doi.org/10.1016/j.neuropsychologia.2009.05.020

Coutanche, M. N., & Thompson-Schill, S. L. (2014). Creating Concepts from Converging Features in Human Cortex. *Cerebral Cortex*, bhu057. http://doi.org/10.1093/cercor/bhu057

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. S. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain*, *126*(5), 1193–1201. http://doi.org/10.1093/brain/awg104

Davey, J., Cornelissen, P., Thompson, H., 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. The Journal of Neuroscience, 35(46), 15230 – 15239

Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., Krieger-Redwood, K., Bernhardt, B. C., Smallwood, J & Jefferies, E. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes. *NeuroImage*, *137*, 165–177.

Devlin, J. T., Matthews, P. M. & Rushworth, M. F. S. (2003), Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive neuroscience, 15:1,* 71-84

Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Moss, H. E., Matthews, P. M & Tyler, L. K. (2000). Susceptibility-Induced Loss of Signal: Comparing PET and fMRI on a Semantic Task. *NeuroImage*, *11*(6), 589–600. http://doi.org/10.1006/nimg.2000.0595

Devlin, J. T., & Watkins, K. E. (2007). Stimulating language: insights from TMS. *Brain*, *130*(3), 610–622. http://doi.org/10.1093/brain/awl331

Dikker, S., & Pylkkänen, L. (2011). Before the N400: Effects of lexical–semantic violations in visual cortex. *Brain and Language*, *118*(1–2), 23–28. http://doi.org/10.1016/j.bandl.2011.02.006

Dikker, S., & Pylkkänen, L. (2013). Predicting Language: MEG Evidence for Lexical Preactivation. *Brain and Language* 127(1), 55–64. doi:10.1016/j.bandl.2012.08.004.

Duecker, F., de Graaf, T. A., Jacobs, C., & Sack, A. T. (2013). Time- and task-dependent non-neural effects of real and sham TMS. *PloS One*, *8*(9), e73813.

Fruchter, J., Linzen, T., Westerlund, M. & Marantz, A. (2015). Lexical Preactivation in Basic Linguistic Phrases. *Journal of Cognitive Neuroscience* 27(10), 1912–35. doi:10.1162/jocn_a_00822.

Fujimaki, N., Hayakawa, T., Ihara, A., Wei, Q., Munetsuna, S., Terazono, Y., Murata, T. (2009). Early neural activation for lexico-semantic access in the left anterior temporal area analyzed by an fMRI-assisted MEG multidipole method. *NeuroImage*, *44*(3), 1093–1102.

Gagnon, G., Schneider, C., Grondin, S., & Blanchet, S. (2011). Enhancement of episodic memory in young and healthy adults: A paired-pulse TMS study on encoding and retrieval performance. *Neuroscience Letters*, *488*(2), 138–142. http://doi.org/10.1016/j.neulet.2010.11.016

Gerwig, M., Kastrup, O., Meyer, B.U., Niehaus, L. (2003). Evaluation of cortical excitability by motor and phosphine thresholds in transcranial magnetic stimulation. *Journal of the Neurological Sciences*, 215, 75-78.

Gold, B., Balota, D., Jones, S., Powell, D., Smith, C., & Andersen, A. (2006). Dissociation of Automatic and Strategic Lexical semantics: Functional Magnetic Resonance Imaging Evidence for Differing Roles of Multiple Frontotemporal Regions. *The Journal of Neuroscience*, *26*(24), 6523–6532.

Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like Magnetoencephalography Responses Modulated by Semantic Context, Word Frequency, and Lexical Class in Sentences. *NeuroImage*, *17*(3), 1101–1116. https://doi.org/10.1006/nimg.2002.1268

Hallam, Glyn P., Carin Whitney, Mark Hymers, Andre D. Gouws, and Elizabeth Jefferies. "Charting the Effects of TMS with fMRI: Modulation of Cortical Recruitment within the Distributed Network Supporting Semantic Control." *Neuropsychologia* 93, Part A (December 2016): 40–52. doi:10.1016/j.neuropsychologia.2016.09.012.

Hall, E. L., Robson, S. E., Morris, P. G., & Brookes, M. J. (2014). The relationship between MEG and fMRI. *NeuroImage*, *102*, *Part 1*, 80–91. http://doi.org/10.1016/j.neuroimage.2013.11.005

Hallam, G.P., Whitney, C., Hymers, M., Gouws, A.D., Jefferies, E. (2016) Charting the effects of TMS with fMRI: Modulation of cortical recruitment within the distributed network supporting semantic control. *Neuropsychologia*, 93, 40-52.

Hanslmayr, S., Staresina, B. P., & Bowman, H. (2016). Oscillations and Episodic Memory: Addressing the Synchronization/Desynchronization Conundrum. *Trends in Neurosciences*, *39*(1), 16–25. http://doi.org/10.1016/j.tins.2015.11.004

Hanslmayr, S., Staudigl, T., & Fellner, M.-C. (2012). Oscillatory power decreases and long-term memory: the information via desynchronization hypothesis. *Frontiers in Human Neuroscience*, *6*. http://doi.org/10.3389/fnhum.2012.00074

Hanslmayr, S., Volberg, G., Wimber, M., Raabe, M., Greenlee, M. W., & Bäuml, K.-H. T. (2011). The Relationship between Brain Oscillations and BOLD Signal during Memory Formation: A Combined EEG–fMRI Study. *The Journal of Neuroscience*, *31*(44), 15674–15680. http://doi.org/10.1523/JNEUROSCI.3140-11.2011

Helenius, P., Salmelin, R. Service, E. & Connolly, J. F. (1998) Distinct Time Courses of Word and Context Comprehension in the Left Temporal Cortex. *Brain* 121(6) 1133–1142

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: 250–66. doi:10.1016/j.cortex.2014.09.001.

Hoffman, P., Jefferies, E., & Lambon Ralph, M. A. (2010). Ventrolateral Prefrontal Cortex Plays an Executive Regulation Role in Comprehension of Abstract Words: Convergent Neuropsychological and Repetitive TMS Evidence. *The Journal of Neuroscience*, *30*(46), 15450–15456. http://doi.org/10.1523/JNEUROSCI.3783-10.2010

Holmes, A. P., Blair, R. C., Watson, J. D. G., & Ford, I. (1996). Nonparametric Analysis of Statistic Images from Functional Mapping Experiments. *Journal of Cerebral Blood Flow & Metabolism*, *16*(1), 7–22. http://doi.org/10.1097/00004647-199601000-00002

Hauk, O., Davis, M.H, Ford, M., Pulvermuller, F., Marslen-Wilson, W.D. (2006) The time-course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, 30, 1383-1400.

Huang, M. X., Mosher, J. C., & Leahy, R. M. (1999). A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Physics in Medicine and Biology*, *44*(2), 423. http://doi.org/10.1088/0031-9155/44/2/010

Huang, M. X., Shih, J. J., Lee, R. R., Harrington, D. L., Thoma, R. J., Weisend, M. P., Hanlon, F., Paulson, K. M., Li, T. & Martin, K., Miller, G. A. & Canive, J. M. (2004). Commonalities and differences among vectorized beamformers in electromagnetic source imaging. *Brain Topography*, *16*(3), 139–158.

Hymers, M., Prendergast, G., Johnson, S. R., & Green, G. G. R. (2010). Source stability index: A novel beamforming based localisation metric. *NeuroImage*, *49*(2), 1385–1397. http://doi.org/10.1016/j.neuroimage.2009.09.055

Ihara, A., Hayakawa, T., Wei, Q., Munetsuna, S., Fujimaki, N. (2007) Lexical access and selection of contextually appropriate meaning for ambiguous words. Neuroimage, 38, 576-588.

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. *Journal of Neuroscience*, *36*(5), 1490–1501. https://doi.org/10.1523/JNEUROSCI.2999-15.2016

Jefferies, E. (2013). The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. *Cortex*, *49*(3), 611–625.

Jefferies, E. & Lambon Ralph, M. A. (2006). Semantic Impairment in Stroke Aphasia Versus Semantic Dementia: a Case-series Comparison. *Brain* 129(8), 2132–2147

JeYoung, J. & Lambon Ralph, M. A. (2016). Mapping the Dynamic Network Interactions Underpinning Cognition: A cTBS-fMRI Study of the Flexible Adaptive Neural System for Semantics." *Cerebral Cortex* 26(8): 3580–90. doi:10.1093/cercor/bhw149.

Kielar, A., Meltzer, J., A., Moreno, S., Alain, C. &Bialystok, E. (2014). Oscillatory Responses to Semantic and Syntactic Violations. *Journal of Cognitive Neuroscience* 26(12), 2840–62. doi:10.1162/jocn_a_00670.

Klein, M., Grainger, J., Wheat, K. L., Millman, R. E., Simpson, M. I. G., Hansen, P. C., & Cornelissen, P. L. (2014). Early Activity in Broca's Area During Reading Reflects Fast Access to Articulatory Codes From Print. *Cerebral Cortex*, bht350. http://doi.org/10.1093/cercor/bht350

Kozinska, D., Carducci, F. & Nowinski, K. (2001). Automatic Alignment of EEG/MEG and MRI Data Sets'. *Clinical Neurophysiology* 112(8), 1553–1561

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*(1), 42–55. https://doi.org/10.1038/nrn.2016.150

Lambon Ralph, M. A., Pobric, G., & Jefferies, E. (2009). Conceptual Knowledge Is Underpinned by the Temporal Pole Bilaterally: Convergent Evidence from rTMS. *Cerebral Cortex*, *19*(4), 832–838. http://doi.org/10.1093/cercor/bhn131

Lau, E., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. (2013). Automatic Semantic Facilitation in Anterior Temporal Cortex Revealed through Multimodal Neuroimaging. *The Journal of Neuroscience*, *33*(43), 17174–17181.

Lau, E. F., Phillips, C. & Poeppel, D. (2008). A Cortical Network for Semantics: (De)constructing the N400. *Nature Reviews Neuroscience* 9(12): 920–33. doi:10.1038/nrn2532.

Lau, E. F., Weber, K., Gramfort, A., Hämäläinen, M. S., & Kuperberg, G. R. (2014). Spatiotemporal Signatures of Lexical–Semantic Prediction. *Cerebral Cortex*, bhu219. http://doi.org/10.1093/cercor/bhu219

Lewis, A. G., & Bastiaansen, M. (2015). A Predictive Coding Framework for Rapid Neural Dynamics during Sentence-Level Language Comprehension.” *Cortex*, 68, 155–68. doi:10.1016/j.cortex.2015.02.014.

Littel, R. C., Stroup, W. W., Milliken, G. A., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for Mixed Models, Second Edition*. SAS Institute.

Luo, Y., Zhang, Y., Feng, X. & Zhou, X. (2010) Electroencephalogram Oscillations Differentiate Semantic and Prosodic Processes during Sentence Reading. *Neuroscience* 169(2), 654–64. doi:10.1016/j.neuroscience.2010.05.032.

Maess, B., Herrmann, C. S., Hahne, A., Nakamura, A., & Friederici, A. D. (2006). Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. *Brain Research*, *1096*(1), 163–172. https://doi.org/10.1016/j.brainres.2006.04.037

Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, *38*(3), 487–497.

Martin, A. (2007). The Representation of Object Concepts in the Brain. *Annual Review of Psychology*, *58*(1), 25–45. http://doi.org/10.1146/annurev.psych.57.102904.190143

Martin, A, Haxby, J.V., Lalonde, F. M., Wiggs, C.L. & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270(5233), 102–105

Mion, M., Patterson, K., Acosta-Cabronero, J., Pengas, G., Izquierdo-Garcia, D., Hong, Y. T., Fryer, T. D., Williams, G. B., Hodges, J. R. & Nestor, P. J. (2010). What the left and right anterior fusiform gyri tell us about semantic memory. *Brain*, 133(11), 3256–3268. http://doi.org/10.1093/brain/awq272

Mollo, G., Cornelissen, P. L., Millman, R. E., Ellis, A. E. & Jefferies, E. (2017) "Oscillatory Dynamics Supporting Semantic Cognition: MEG Evidence for the Contribution of the Anterior Temporal Lobe Hub and Modality-Specific Spokes." *PLOS ONE* 12(1), e0169269. doi:10.1371/journal.pone.0169269.

Mollo, G., Jefferies, E., Cornelissen, P., Gennari, S.P. (2018 in press) Context-dependent lexical ambiguity resolution: MEG evidence for the time-course of activity in inferior frontal gyrus and posterior middle temporal gyrus. *Brain and Language*.

Murphy, C., Rueschemeyer, S-A., Watson, D., Karapanagiotidis, T., Smallwood, J. & Jefferies, E. (2017). Fractionating the Anterior Temporal Lobe: MVPA Reveals Differential Responses to Input and Conceptual Modality. *NeuroImage* 147: 19–31. doi:10.1016/j.neuroimage.2016.11.067.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology*, *106*(3), 226–254.

Neely, J. (1991) Semantic priming effects in visual word recognition: a selective review of current findings and theories. In Besner, D., & Humphreys, G. W. (Ed.), *Basic Processes in Reading: Visual Word Recognition*. Psychology Press.

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 Prefrontal and Temporo-parietal Cortices. *Journal of Cognitive Neuroscience*, *22*(7), 1597–1613. http://doi.org/10.1162/jocn.2009.21289

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. http://doi.org/10.1162/jocn_a_00442

Noppeney, U., Phillips, J., & Price, C. (2004). The neural areas that control the retrieval and selection of semantics. *Neuropsychologia*, *42*(9), 1269–1280. http://doi.org/10.1016/j.neuropsychologia.2003.12.014

Noppeney, U. & Price, C.J. (2002). A PET Study of Stimulus- and Task-Induced Semantic Processing’. *NeuroImage* 15(4), 927–935

Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience – virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology*, *10*(2), 232–237. http://doi.org/10.1016/S0959-4388(00)00081-7

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.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2007). Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. Proceedings of the National Academy of Sciences, 104(50), 20137-20141.

Pobric, G., Lambon Ralph, M. A., & Jefferies, E. (2009). The role of the anterior temporal lobes in the comprehension of concrete and abstract words: rTMS evidence. *Cortex*, *45*(9), 1104–1110.

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Category-Specific versus Category-General Semantic Impairment Induced by Transcranial Magnetic Stimulation. *Current Biology*, *20*(10), 964–968. http://doi.org/10.1016/j.cub.2010.03.070

Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Amodal semantic representations depend on both anterior temporal lobes: Evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia*, *48*(5), 1336–1342. http://doi.org/10.1016/j.neuropsychologia.2009.12.036

Prendergast, G., Johnson, S. R., Hymers, M., Woods, W., & Green, G. G. R. (2011). Non-parametric statistical thresholding of baseline free MEG beamformer images. *NeuroImage*, *54*(2), 906–918. http://doi.org/10.1016/j.neuroimage.2010.08.005

Rice, G. E., Lambon Ralph, M. A., & Hoffman, P. (2015). The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cerebral Cortex*, bhv024. http://doi.org/10.1093/cercor/bhv024

Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., & Price, C. J. (2006). Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, & Behavioral Neuroscience*, *6*(3), 201–213.

Rorden, C., Karnath, H.-O., & Bonilha, L. (2007). Improving lesion-symptom mapping. *Journal of Cognitive Neuroscience*, *19*(7), 1081–1088. http://doi.org/10.1162/jocn.2007.19.7.1081

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, *120*(12), 2008–2039. http://doi.org/10.1016/j.clinph.2009.08.016

Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400–2406. http://doi.org/10.1093/brain/123.12.2400

Sereno, S.C., Brewer, C.C., O'Donnell, P.J. (2003) Context effects in word recognition: Evidence for early interactive processing. Psychological Science, 14, 328-333.

Singh, K. D., Barnes, G. R., Hillebrand, A., Forde, E. M. E., & Williams, A. L. (2002). Task-Related Changes in Cortical Synchronization Are Spatially Coincident with the Hemodynamic Response. *NeuroImage*, *16*(1), 103–114. http://doi.org/10.1006/nimg.2001.1050

Sliwinska, M. W. W., Khadilkar, M., Campbell-Ratcliffe, J., Quevenco, F., & Devlin, J. T. (2012). Early and sustained supramarginal gyrus contributions to phonological processing. *Language Sciences*, *3*, 161.

Smallwood, J., Karapanagiotidis, T., Ruby, F., Medea, B., Caso, I. de, Konishi, M., Wang, H.-T., Hallam, G., Marguiles, D. S. & 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. https://doi.org/10.1371/journal.pone.0152272

Stockwell, R. G., Mansinha, L., & Lowe, R. P. (1996). Localization of the complex spectrum: the S transform. *IEEE Transactions on Signal Processing*, *44*(4), 998–1001. http://doi.org/10.1109/78.492555

Strafella, A. P., & Paus, T. (2001). Cerebral Blood-Flow Changes Induced by Paired-Pulse Transcranial Magnetic Stimulation of the Primary Motor Cortex. *Journal of Neurophysiology*, *85*(6), 2624–2629.

Thompson, H. E., Robson, H., Lambon Ralph, M. A. & Jefferies, E. (2015). Varieties of Semantic 'access' Deficit in Wernicke's Aphasia and Semantic Aphasia. *Brain* 138(12), 3776–92. doi:10.1093/brain/awv281.

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* 94(26), 14792–14797

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P. & Knight, R. T. (1998). Verb Generation in Patients with Focal Frontal Lesions: A Neuropsychological Test of Neuroimaging Findings. *Proceedings of the National Academy of Sciences* 95(26), 15855–60. doi:10.1073/pnas.95.26.15855.

Urooj, U., Cornelissen, P. L., Simpson, M. I. G., Wheat, K. L., Woods, W., Barca, L., & Ellis, A. W. (2014). Interactions between visual and semantic processing during object recognition revealed by modulatory effects of age of acquisition. *NeuroImage*, *87*, 252–264. http://doi.org/10.1016/j.neuroimage.2013.10.058

Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*(6597), 254–256.

Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *Biomedical Engineering, IEEE Transactions on*, *44*(9), 867–880.

Visser, M., Jefferies, E., Embleton, K. V., & Lambon Ralph, M. A. (2012). Both the Middle Temporal Gyrus and the Ventral Anterior Temporal Area Are Crucial for Multimodal Semantic Processing: Distortion-corrected fMRI Evidence for a Double Gradient of Information Convergence in the Temporal Lobes. *Journal of Cognitive Neuroscience*, *24*(8), 1766–1778.

Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic Processing in the Anterior Temporal Lobes: A Meta-analysis of the Functional Neuroimaging Literature. *Journal of Cognitive Neuroscience*, *22*(6), 1083–1094. http://doi.org/10.1162/jocn.2009.21309

Vitello, S., Warren, J. E., Devlin, J. T. & Rodd, J. M. (2014).Roles of Frontal and Temporal Regions in Reinterpreting Semantically Ambiguous Sentences. *Frontiers in Human Neuroscience* 8: 1-14. doi:10.3389/fnhum.2014.00530.

Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience*, *1*(1), 73–80. http://doi.org/10.1038/35036239

Wang, L., Jensen, O., van den Brink, D., Weder, N., Schoffelen, J-M., Magyari, L., Hagoort, P. &Bastiaansen, M. (2012). Beta Oscillations Relate to the N400m during Language Comprehension. *Human Brain Mapping* 33(12), 2898–2912. doi:10.1002/hbm.21410.

Westerlund, M., & Pylkkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia*, *57*, 59–70. http://doi.org/10.1016/j.neuropsychologia.2014.03.001

Wheat, K. L., Cornelissen, P. L., Frost, S. J., & Hansen, P. C. (2010). During Visual Word Recognition, Phonology Is Accessed within 100 ms and May Be Mediated by a Speech Production Code: Evidence from Magnetoencephalography. *Journal of Neuroscience*, *30*(15), 5229–5233. http://doi.org/10.1523/JNEUROSCI.4448-09.2010

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. *Cerebral Cortex*, *21*(4), 831–844. http://doi.org/10.1093/cercor/bhq148

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2011). Executive Semantic Processing Is Underpinned by a Large-scale Neural Network: Revealing the Contribution of Left Prefrontal, Posterior Temporal, and Parietal Cortex to Controlled Retrieval and Selection Using TMS. *Journal of Cognitive Neuroscience*, *24*(1), 133–147. http://doi.org/10.1162/jocn_a_00123

Wirth, M., Jann, K., Dierks, T., Federspiel, A., Wiest, R., & Horn, H. (2011). Semantic memory involvement in the default mode network: A functional neuroimaging study using independent component analysis. *NeuroImage*, *54*(4), 3057–3066. https://doi.org/10.1016/j.neuroimage.2010.10.039

Woodhead, Z.V., Barnes, G.R., Penny, W., Moran, R., Teki, S., Price, C.J., Leff, A.P. (2014) Reading front to back: MEG evidence for early feedback effects during word recognition. *Cerebral Cortex*, 24, 817-25.

Xu, Y., Qixiang, L., Zaizhu, H., Yong, H. & Yanchao, B. (2016). Intrinsic Functional Network Architecture of Human Semantic Processing: Modules and Hubs. *NeuroImage* 132, 542–55. doi:10.1016/j.neuroimage.2016.03.004.

Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari., D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of Neurophysiology, 106, 1125-1165.

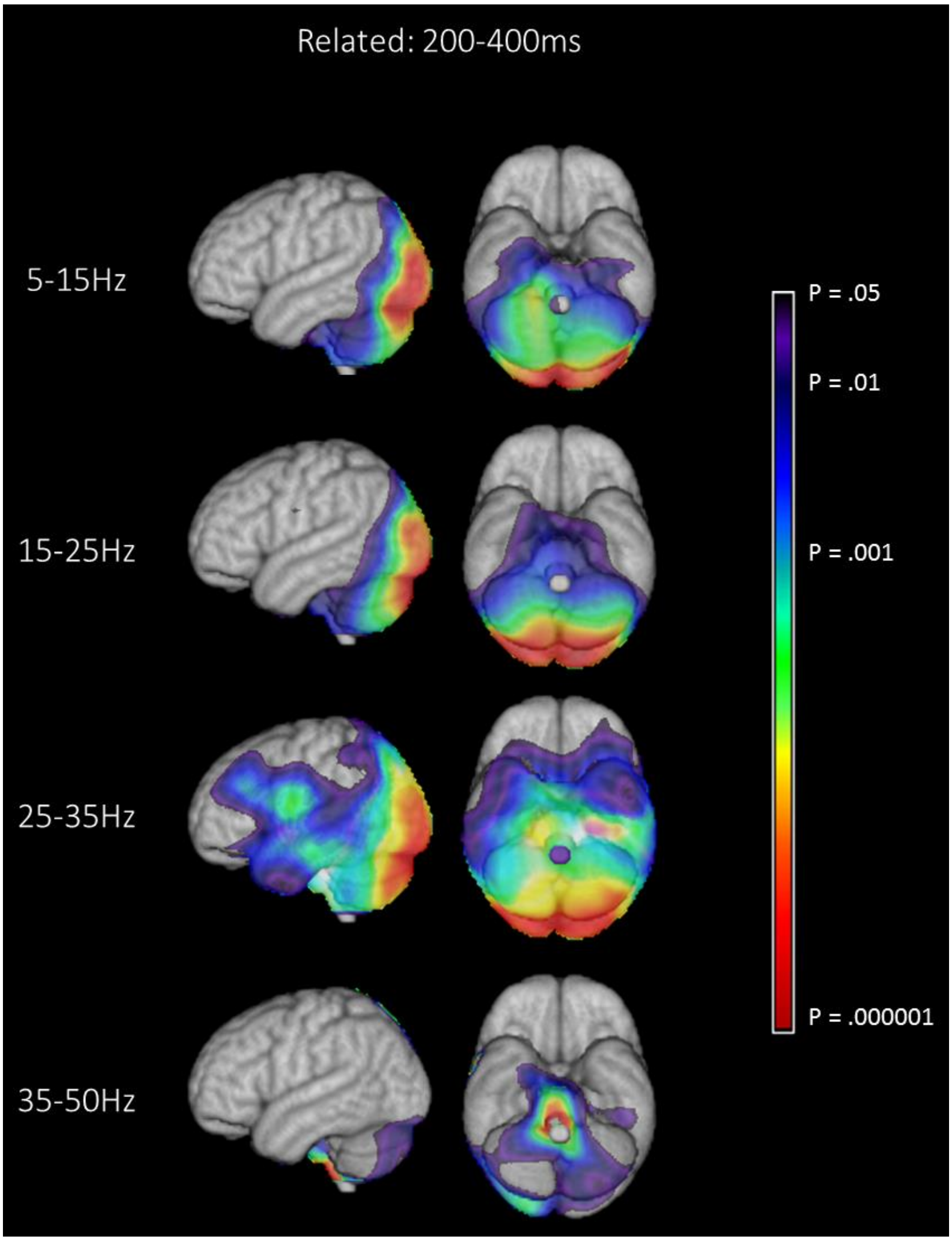
Yvert, G., Perrone-Bertolotti, M., Baciu, M., & David, O. (2012). Dynamic Causal Modeling of Spatiotemporal Integration of Phonological and Semantic Processes: An Electroencephalographic Study. *Journal of Neuroscience*, *32*(12), 4297–4306. http://doi.org/10.1523/JNEUROSCI.6434-11.2012

Zhang, L., & Pylkkänen, L. (2015). The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *NeuroImage*, *111*, 228–240. http://doi.org/10.1016/j.neuroimage.2015.02.028

Supplementary Materials

Supplementary Analysis 1: Whole-brain beamforming for a range of frequency bands

The most extensive changes in total power in response to the task were power decreases in the 25-35Hz frequency band. Data for all frequency bands from 200-400ms are provided in Supplementary Figure 1 below.

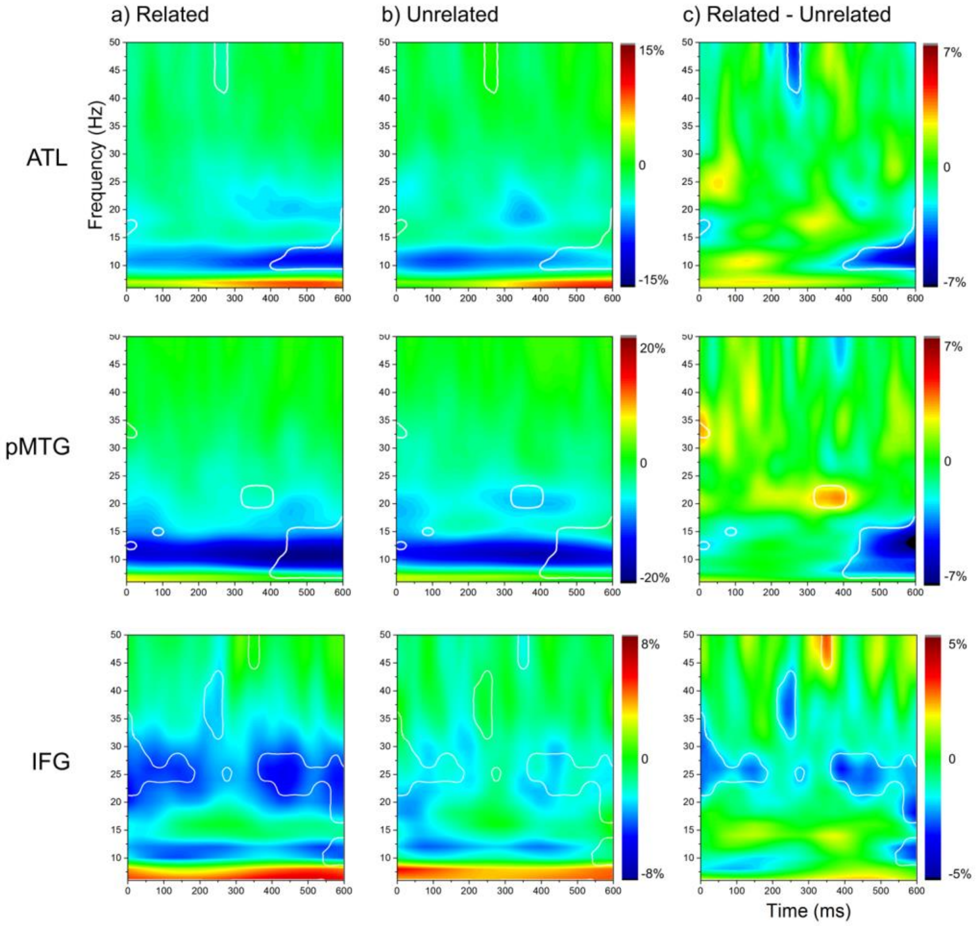


Supplementary Figure 1: Additional frequency bands for the 200-400ms time window, related condition

Supplementary Analysis 2: Contrasts between related and unrelated words within the temporal lobe

In addition to the strong/weak differences focused on in the main analysis, we also computed differences between related and unrelated trials to allow comparison with previous studies that employed similar contrasts (for a review, see Lau et al., 2008). The results can be seen in Supplementary Figure 2. In line with this literature, ATL, pMTG and IFG all showed stronger task-related changes in oscillatory power for semantically-associated compared with unrelated items, consistent with the purported role of these sites in semantic retrieval. At ATL and pMTG, there was more sustained engagement for related items at around 10Hz, producing a marked difference between conditions around 400ms post-onset. ATL also showed an earlier response to semantic relatedness around 250ms, from 40-50Hz, demonstrating a relatively rapid sensitivity to semantic variables in ATL, even though the effect of semantic coherence (strong > weak) emerged later, consistent with previous findings (Lau et al., 2013). In pMTG, oscillatory power was also greater for unrelated than related words at 350-400ms at 20Hz. Effects of relatedness were most marked in IFG, which showed both early and late effects, from 0-200 and 400-600ms, centred on 25Hz.

Discussion: The effects of relatedness and strength of association were similar in time-frequency space in ATL – i.e., both effects occurred from 400ms and from 10-15Hz. These effects might reflect a building pattern of coherent semantic retrieval, which is strengthened when related items are highly associated. In pMTG, however, the stronger response to weak associations started soon after the onset of the second word and lasted until 400ms, when the stronger response to related items commenced. If pMTG plays a role in representing a semantic context which can differ from the pattern of long-term associations in ATL, the early and sustained response from 50-400ms could reflect efforts to identify a context in which two weakly-related (or potentially unrelated) words can be coherently associated (while for strong association trials, an appropriate context is already present at the onset of the second word). Finally, in IFG, the early effect of relatedness overlapped with the weak>strong effect at 50ms and 25Hz, suggesting this initial response to meaning reflected controlled retrieval processes. Later, however, IFG showed both effects of relatedness and strong>weak associations: presentation of strong associations expanded the response to related items to include lower frequencies (15Hz to 20Hz. Therefore, we speculate that by 400ms, a pattern of coherent retrieval across semantic sites might be established, eliciting an overall effect of relatedness at all three sites.



Supplementary Figure 2: a) Percentage signal change in the related condition, relative to baseline. b) Percentage signal change in the unrelated condition, relative to baseline. c) Percentage signal change between related and unrelated conditions, separately for ATL and pMTG. White lines are derived from the statistical comparison between related and unrelated conditions. The boundaries enclose regions fulfilling two criteria: i) percentage signal change between the related and unrelated conditions is significantly different from zero (p<0.05) and ii) percentage signal change computed separately for each condition is significantly different from zero for at least one of the two conditions.

Yellow-red colours indicate regions of *power increase* relative to the baseline, while cyan-blue indicates *power decreases* relative to the baseline, and green indicates no change from baseline.

Supplementary Analysis 3: Additional TMS analyses

The supplementary analyses below examine (i) the effects of TMS on accuracy for strong and weak associations, and (ii) the effects of TMS on the semantic and number tasks overall. In this analysis, there are 14 as opposed to 15 participants, since data for the even/odd digit task were not recorded for one participant. The results of Analysis 1 motivated the inclusion of accuracy as a covariate in the analysis of response time above. Results from Analysis 2 reveal non-specific effects of TMS and therefore motivated the inclusion of the number control task as a covariate in the main analyses reported above. In addition, Supplementary Table 1 provides full summary statistics for this experiment.

Supplementary Table 1: Behavioural data for Experiment 2 (Chronometric TMS)

Measurement	Site	Condition	0 & 40ms	125 & 165ms	250 & 290ms	450 & 490ms	No TMS	0 & 40ms	125 & 165ms	250 & 290ms	450 & 490ms	No TMS
			Mean					SD				
RT	ATL	strong	539.53	559.05	584.54	616.85	573.43	120.50	138.41	156.16	195.46	143.09
	ATL	weak	601.16	613.62	641.42	704.06	630.98	167.37	174.36	192.62	214.26	163.98
	ATL	number	535.45	565.54	576.62	616.55	571.87	144.29	163.49	187.26	210.41	165.70
	pMTG	strong	616.55	506.63	527.08	516.72	534.54	210.41	132.87	145.15	130.08	119.15
	pMTG	weak	498.46	512.79	529.29	543.91	595.16	105.49	100.93	116.69	141.97	123.20
	pMTG	number	578.90	590.26	585.17	610.58	531.61	156.50	142.16	139.25	164.37	129.47
	SHAM	strong	510.81	499.79	531.80	550.76	551.33	113.94	119.35	119.68	147.79	130.21
	SHAM	weak	559.64	573.55	588.73	617.25	606.62	133.06	144.95	145.19	158.97	125.16
	SHAM	number	492.95	512.32	513.29	524.02	517.83	126.25	132.99	127.71	145.03	127.27
Accuracy	ATL	strong	0.96	0.95	0.96	0.96	0.94	0.04	0.06	0.04	0.05	0.05
	ATL	weak	0.82	0.76	0.83	0.78	0.83	0.13	0.17	0.11	0.18	0.08
	ATL	number	0.92	0.92	0.95	0.95	0.92	0.05	0.05	0.06	0.05	0.04
	pMTG	strong	0.97	0.97	0.97	0.97	0.95	0.05	0.03	0.04	0.03	0.06
	pMTG	weak	0.84	0.83	0.82	0.80	0.85	0.09	0.15	0.11	0.09	0.10
	pMTG	number	0.89	0.94	0.93	0.94	0.93	0.12	0.04	0.06	0.06	0.04
	SHAM	strong	0.97	0.97	0.97	0.97	0.95	0.04	0.04	0.04	0.04	0.06
	SHAM	weak	0.82	0.78	0.78	0.77	0.85	0.11	0.09	0.15	0.14	0.11
	SHAM	number	0.95	0.95	0.96	0.94	0.94	0.07	0.04	0.05	0.05	0.05

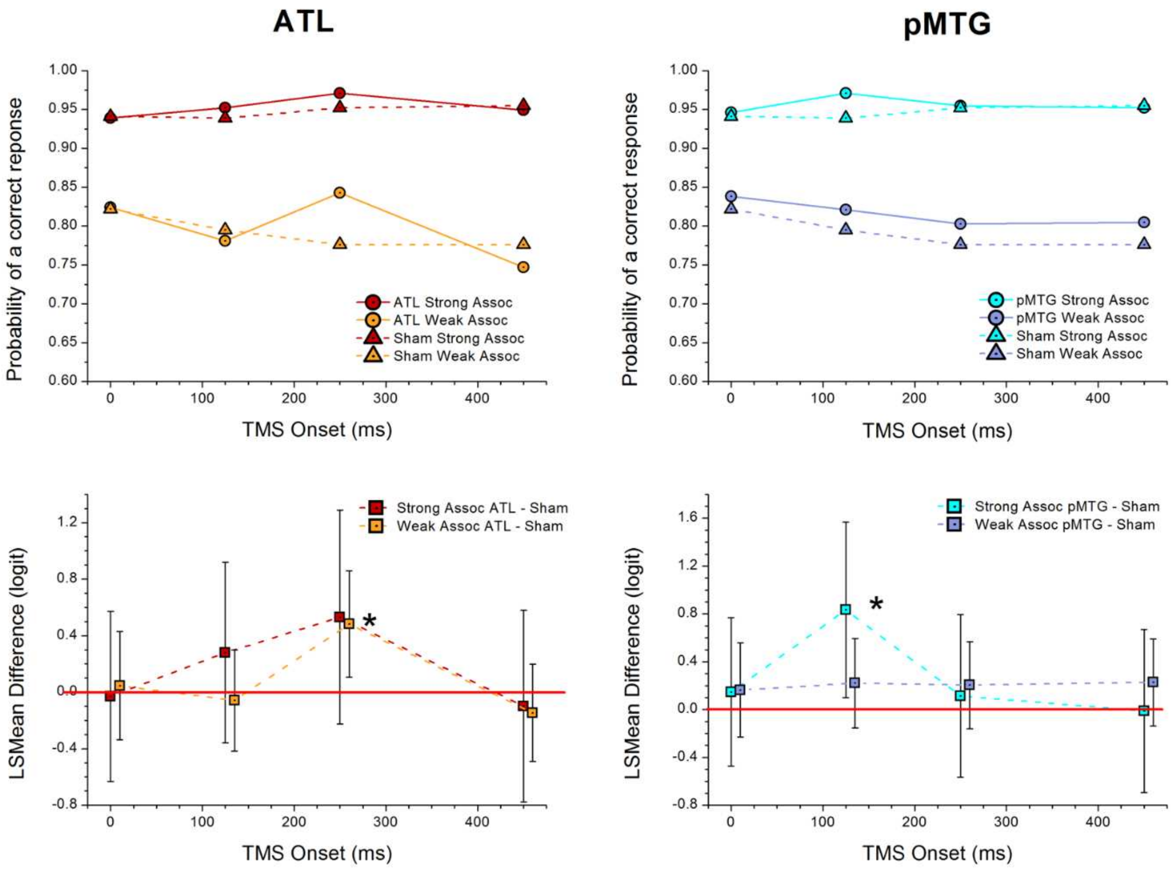
Effect of TMS on accuracy for strong and weak associations

We wanted to check whether the accuracy of speeded judgements about strong and weak semantic relationships between pairs of words was affected by TMS. To test this, we pooled the total number of correct responses, or hits, for the total number of trials in each block, separately for participant, stimulation site, TMS timing, condition and session. We used these pooled data as input to a generalised linear mixed model (GLMM), implemented in PROC GLIMMIX in SAS v9.4. Since accuracy data have a binomial distribution (hit = 1, 0 = error), we used a logistic link function in the model for the outcome. For the starting model, the fixed effects were: site (ATL, pMTG, sham), TMS time, condition (strong vs. weak) and their two- and three-way interactions. In addition, we included block number and session number as covariates and the per-subject intercept as a random effect. The criteria we used to optimize the final model were: (i) a significant reduction in -2Log-Likelihood between it and the empty model, (ii) only explanatory variables that were statistically significant at p<.05 should be retained. The exception to this was that we had to include the three way interaction between condition (strong vs. weak association), TMS time and site in order to compute post-hoc pairwise comparisons. Supplementary Table 2 shows the outcome for the fixed effects in the final model for which the generalised chi-square per degree of freedom was 1.05, suggesting a good model fit with no over-dispersion. Once the final model was fitted, we used PROC GLIMMIX to estimate pairwise t-tests to compare the LS mean accuracy with and without TMS, carried out separately at each site for each condition. These post-hoc comparisons were controlled for multiple comparisons.

Supplementary Table 2: Effect of TMS on accuracy for strong and weak associations

Model Parameter	F-value (DF)	p-value	-2Log likelihood
Empty model			63784.3
Condition	405.40 (1, 1399)	<.001	63277.5
Condition × Time	2.40 (6, 1399)	.026	
Condition × Time × Site	1.17 (16, 1399)	.29	
Block order	3.92 (3,1399)	.0084	

Plotted separately for ATL and pMTG, Supplementary Figure 3 (upper row) shows the proportion of correct responses for the strong and weak conditions for both the experimental and the sham data. Supplementary Figure 3 (lower row) shows post-hoc paired t-test comparisons between the experimental and sham data, separately for the strong and weak conditions. We found a statistically significant difference between sham and ATL stimulation for weak associations at ~250ms (t(1399) = 2.37, p = 0.018 and between sham and pMTG stimulation for strong associations at ~125ms (t(1399) = 2.08, p = 0.038). These results reflect facilitation of performance, potentially following increases in alertness or motivation (Devlin & Watkins, 2007).



Supplementary Figure 3: Effect of TMS on accuracy for strong and weak associations; TOP ROW: Accuracy for strong and weak association for ATL (left) and pMTG (right) relative to sham stimulation. BOTTOM ROW: LS Means difference in accuracy for ATL (left) and pMTG (right) relative to sham site. Points above the red line indicates higher accuracy relative to sham, whereas point below the red line indicates lower accuracy relative to sham. Differences in accuracy (which in all cases correspond to TMS-induced behavioural facilitation) that were statistically significant at p<.05 are indicated by asterisks. Error bars show 95% confidence intervals.

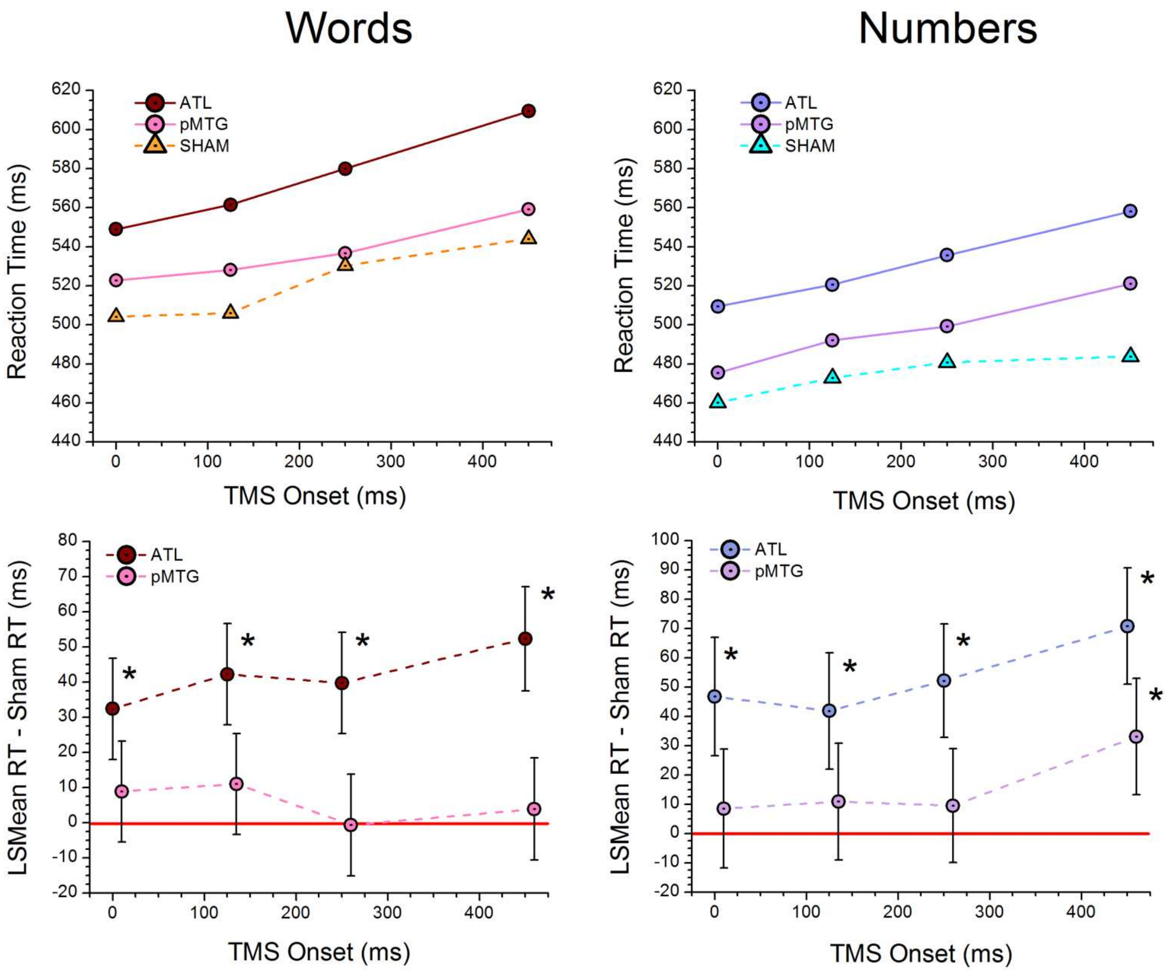
Effect of TMS on response times in the semantic word vs. number control task

We wanted to test whether there might be non-specific effects of administering chronometric TMS that should be controlled for in the main analysis of reaction times in relation to strong versus weal semantic association. To do this we compared reaction times in the digit parity task with the semantic task overall. We used PROC MIXED in SAS v9.4 to build an initial GLMM with the following fixed effects: site (ATL, pMTG, sham), TMS time, condition (semantic vs. number task) and their two- and three-way interactions. The starting model also included block number, session number and mean accuracy per block as covariates. We included the per-subject intercept as a random effect. Incorrect responses and outlying data points that fell more than 2 SD from each participant's mean RT were removed, for each session, prior to analysis. The criteria we used to optimize the final model were: (i) a significant reduction in -2Log-Likelihood between it and the empty model, (ii) only explanatory variables that were statistically significant at p<.05 should be retained. Once the final model was fitted, we used PROC MIXED to estimate pairwise t-test comparisons between experiment and sham reaction times at each time point for TMS administration, computed separately for ATL and pMTG as well as word versus number stimuli. All post-hoc comparisons were controlled for multiple comparisons. The fixed effects that were retained in the optimized model are reported in Supplementary Table 3. We found statistically significant main effects of TMS time, condition (number vs. word task) and site (ATL vs. pMTG vs. sham), and we retained only the non-significant three-way interaction in order that we could compute the pair-wise post-hoc comparisons. The covariates of block and session order were also statistically significant at p<.05.

Supplementary Table 3: Effect of TMS on RT to semantic and number parity tasks

Model Parameter	F-value (DF)	Z-value	p-value	-2Log likelihood
Empty model				146110.3
Time	53.80 (3, 11000)		<.001	144796.6
Site	122.78 (2, 11000)		<.001	
Condition × Time × Site	1.20 (17, 11000)		.26	
Block order	41.10 (3, 11000)		<.001	
Testing session	178.80 (2, 11000)		<.001	
Participant covariance		2.74	.0031	

As illustrated in Supplementary Figure 4, post-hoc comparisons for the semantic task showed statistically significant differences between LS mean RT for ATL stimulation compared to sham at all four time points (0ms: t(11000) = 4.38, p <.001; 125ms: t(11000) = 5.72, p <.001; 250ms: t(11000) = 5.38, p <.001; 450ms: t(11000) = 6.90, p <.001). We found no equivalent differences between sham and pMTG stimulation. Post-hoc comparisons for the number task showed statistically significant differences in LS mean RT for the ATL stimulation compared to sham at all four time points (0ms: t(11000) = 4.50, p <.001; 125ms: t(11000) = 4.10, p <.001; 250ms: t(11000) = 5.25, p <.001; 450ms: t(11000) = 6.97, p <.001), as well as for pMTG relative to sham at 450ms (t(11000) = 3.25, p = 0.0012). These results suggest that despite the use of a staircase procedure (Sliwinska et al., 2012) which was designed to reduce sensitivity to the variability in TMS onset time (see Methods), RTs increased with systematically with TMS onset time. This may correspond to an expectancy effect in participants who waited for the TMS pulse to have occurred before responding. This interpretation is made all the more plausible because we found a similar pattern of increasing RTs following sham stimulation which was administered at 30% of stimulator output. This is not thought to be strong enough to stimulate cortex (Duecker et al., 2013) but is sufficient to produce scalp sensations. The increase in RT, relative to sham, was more marked for ATL than for pMTG, which may reflect the stronger scalp sensations associated with ATL stimulation.



Supplementary Figure 5: Effect of TMS on RT for semantic and digit parity tasks; TOP ROW: RT (ms) for three sites for semantic (left) and digit task (right). BOTTOM ROW: LS Means difference in RT for ATL and pMTG relative to sham site for semantic (left) and digit (right) task. Points above the red line indicate longer RT relative to sham, whereas points below the red line indicate faster RT relative to sham. Differences which are statistically significant at $p<.05$ are indicated by asterisks. Error bars show 95% confidence intervals.