The neural bases of taxonomic and thematic conceptual relations: An MEG study
Gwyneth A. Lewis*, David Poeppel, Gregory L. Murphy

Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, United States

Article history:
Received 2 September 2014
Received in revised form 7 January 2015
Accepted 8 January 2015
Available online 9 January 2015

Keywords:
Taxonomic concepts
Thematic concepts
MEG
Semantic memory
Anterior temporal lobe
Temporoparietal junction

Abstract
Converging evidence from behavioral and neuroimaging studies of human concepts indicate distinct neural systems for taxonomic and thematic knowledge. A recent study of naming in aphasia found involvement of the anterior temporal lobe (ATL) during taxonomic (feature-based) processing, and involvement of the temporoparietal junction (TPJ) during thematic (function-based) processing. We conducted an online magnetoencephalography (MEG) study to examine the spatio-temporal nature of taxonomic and thematic relations. We measured participants’ brain responses to words preceded by either a taxonomically or thematically related item (e.g., cottage—castle, king—castle). In a separate experiment we collected relatedness ratings of the word pairs from participants. We examined effects of relatedness and relation type on activation in ATL and TPJ regions of interest (ROIs) using permutation t-tests to identify differences in ROI activation between conditions as well as single-trial correlational analyses to examine the millisecond-by-millisecond influence of the stimulus variables on the ROIs. Taxonomic relations strongly predicted ATL activation, and both kinds of relations influenced the TPJ. Our results further strengthen the view of the ATL’s importance to taxonomic knowledge. Moreover, they provide a nuanced view of thematic relations as involving taxonomic knowledge.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Taxonomic vs. thematic concepts

Since at least Inhelder and Piaget (1964), taxonomic knowledge has been a major focus of the study of human concepts. Many concepts can be structured into taxonomies, in which specific, concrete categories nest within more general superordinate categories: for example, schnauzer–dog–mammal–vertebrate, dining room table–table–furniture–artifact–object, birthday party–party-social event. Categories that fall within a common superordinate (e.g., all parties) tend to share properties. Furthermore, the nested quality of many such concepts allows one to draw inferences, such as assuming that schnauzers breathe and give birth to live young, even if one has never encountered a schnauzer. Taxonomic categories are generally similarity-based, that is, they have shared attributes. Dogs tend to have four legs, bark, have fur, be pets, and eat meat.

For many years, cognitive psychologists considered taxonomic concepts “real” concepts and other forms of grouping to be the result of immature conceptual systems. For example, children form groupings such as putting a woman with a car because the woman drives the car (e.g., Smiley and Brown, 1979). Such thematic categories are not based on similarity (i.e., shared features) but on extrinsic relations between two objects. Later research discovered that adults also form thematic categories if the task is structured correctly (Estes et al., 2011; Lin and Murphy, 2001; Murphy, 2001). Some adult subjects even prefer thematic to taxonomic categories when forced to make a choice.

Markman (1989) made the important observation that thematic relations are not just a primitive form of concept but are an important part of conceptual knowledge. If you want to know what goes on top of a birthday cake, it does not do any good to know the features of cakes or desserts in general – you have to know that candles go with cake. If you see candles, cake, and balloons, you can infer that a birthday party is taking place. Such relations comprise an important part of our knowledge of events and situations but are theoretically separate from taxonomic knowledge in that taxonomic knowledge tells us the properties of a set of objects, whereas thematic knowledge tells us how other categories relate to that set (Murphy, 2010).

Theoretically, it is unclear whether we store thematic information as part of the same neural network as taxonomic categories. On the one hand, thematic categories have a very different basis. Dogs and leashes may go together (and some people classify them as the same kind of thing), but they do not share properties.
If taxonomic categories are represented by their associations to features (Rogers and McClelland, 2004), this would not seem to include the extrinsic relations thematic categories are based on. On the other hand, the constituents of thematic relations are taxonomic concepts: It is the taxonomic concepts of dogs and leashes that occur in the dog-and-leash thematic concept, so it would not be surprising if the thematic knowledge were mixed in with the taxonomic knowledge. That is, the concept of dog could be linked to a node representing four legs and to another node representing walking on a leash.

Alternatively, thematic knowledge may be part of a different knowledge system, namely, our knowledge of events and situations. At a birthday party one lights candles on a cake, sings “Happy Birthday,” gives presents, and so on, in a particular order. Thus, links between candles and cake may not be part of the representation of those taxonomic categories per se but could instead be embedded in our event knowledge. Studies of the neural representation of these categories could help to distinguish these possibilities.

1.2. Neural responses in taxonomic and thematic associations

The nature of concepts and their neural representation and processing is often investigated in the context of semantic categorization tasks, in which participants determine whether a word falls into a particular category or which of two words goes with a target word. Recent fMRI experiments provide somewhat conflicting results regarding the brain regions involved in thematic and taxonomic categorization. In one study, participants viewed a target and two choice pictures and selected the one picture most similar to the target word. Recent fMRI experiments provide somewhat conflicting results regarding the brain regions involved in thematic and taxonomic categorization. In one study, participants viewed a target and two choice pictures and selected the one picture most related to the target (Kalénine et al., 2009). Stimuli included manipulable and non-manipulable artifacts and natural objects. Half of the related items were related taxonomically, half thematically. They found taxonomic processing involved bilateral occipital regions (especially for non-manipulable natural objects like animals) whereas thematic processing involved temporal and parietal (visuo-motor) regions (especially for manipulable artificial objects like tools). Kalénine et al. concluded that taxonomic relations probably rely on perceptual processes while thematic relations rely on event/action processing, perhaps related to object manipulation. Because subjects viewed three pictures during the task, activation in visual areas is expected. The question remains whether such activation reflects conceptual processes or the perception of pictorial similarity. The Kalénine et al. study offers valuable information about conceptual processing of picture naming, but it is unclear whether such occipital activation would be found without pictures – that is, whether taxonomic knowledge is primarily visual or whether such activation mainly reflects the visual similarity of taxonomically related pictures. Additionally, Carlson et al. (2013) did not find occipital activation correlating with semantic properties when viewing single pictures.

Sachs et al. (2008) attempted to identify neural correlates of taxonomic and thematic relations using a similar choice task with words. In a “biased” condition, a word appeared, followed by an unrelated word appearing with either a taxonomically or thematically related word. Subjects chose the related option. In contrast to Kalénine et al. (2009), Sachs et al. found similar recruitment of occipital, inferior frontal, and middle temporal brain regions for both relations.

An important region for representing conceptual knowledge is the anterior temporal lobe (ATL) (e.g., Rogers et al., 2004), although it is not entirely clear whether its role is restricted to taxonomic knowledge. The ATL is an important region in the neuropathology of semantic dementia (SD), although the full clinical picture includes areas outside the ATL (Gainotti, 2011). Bozeat et al. (2000) measured the performance of SD patients on a semantic decision task (Camel and Cactus Task). In both word and picture versions of the task, participants matched a target (e.g., camel) to a thematically related item (e.g., cactus) from a set of four same-category items (e.g., cactus, sunflower, tree, rose). They found that the performance of SD patients on this task fell well below that of normals and also corresponded with the degree of dementia. The role of the ATL in thematic processing has also been studied in the context of object use. Hodges et al. (2000) found that the ability of SD patients to correctly use an object (e.g., lighting a match) declined as a function of their impairment (e.g., one patient puffed on a match as one would a cigarette). However, it is possible that such effects could arise from loss of taxonomic knowledge, for example, loss of semantic properties of the camel concept. If one forgets what a camel is, thematic knowledge related to it will not be accessible. These studies did not test taxonomic knowledge per se.

There is increasing evidence of the ATL’s importance for taxonomic knowledge. A recent, large aphasia study by Schwartz et al. (2011) specifically contrasted taxonomic and thematic relations. They analyzed a database of naming errors in 86 patients with various lesion locations to identify which locations corresponded with different error types. When shown a picture of a dog, aphasics with naming problems might give the name of another animal (e.g., cat), which would be a taxonomic error. Less often, a thematic error might be produced (e.g., leash or bone). To isolate the two error types, Schwartz et al. regressed out the mutual variance of taxonomic and thematic errors. Their results suggest that two left-hemisphere areas predict such errors: ATL lesions uniquely predicted taxonomic errors whereas temporoparietal junction (TPJ) lesions uniquely predicted thematic errors. Jefferies and Lambon Ralph (2006) also reported naming errors for patients with either semantic dementia resulting from ATL degeneration or aphasia resulting from temporal or frontal damage (or both). They reported that the latter group’s errors were “associative” responses 27% of the time, such as squirrel → nuts, glass → ice, and lorry → diesel. These are thematic relations. The ATL group’s errors were such responses only 1% of the time and were taxonomic responses the rest of the time.

Results from an eye-tracking experiment further support a distinction between the representation of taxonomic and thematic knowledge (Mirman and Graziano, 2012). Aphasics with predominantly ATL or posterior lesions heard individual spoken words, each followed by a screen presenting four images: a target item, a taxonomically or thematically related item, and two distractors. Taxonomic trials yielded longer fixations in the ATL patients, whereas thematic trials led to reduced, delayed fixations in the posterior lesion patients. Taken together, the data from Kalénine et al. (2009), Schwartz et al. (2011), and Mirman and Graziano (2012) suggest different neural implementations of taxonomic and thematic knowledge.

According to a prominent distributed-only model of conceptual knowledge, semantic knowledge is represented in a distributed fashion in modality-specific sensory-motor brain regions (Martin, 2007; Martin and Chao, 2001). As an example, our knowledge of dog includes its typical shape, the sound of its bark, and its gait. Such features may be represented in cortical areas involved in vision, audition, and motion processing, respectively. An alternative view, known as the distributed-plus-hub (AKA spoke-and-hub) model, agrees that specific features of conceptual representations are stored near perceptual or motor areas but proposes that semantic knowledge requires a general, supra-modal mechanism that can generalize across similar concepts differing in some featural way (e.g., poodles and beagles). The ATL may be the “hub” that assimilates features from modality specific regions (the spokes). As evidence, while injury to one of the “spokes” typically results in deficiency of one type of knowledge, damage to the ATL...
Results from previous MEG studies suggest spatially and temporally distinct neural stages in visual word recognition. Upon word presentation, these stages begin with orthographic feature detection in occipital regions at around 100 ms, followed by morphological decomposition in inferior temporal regions at around 150 ms, and retrieval of lexical information in the superior temporal regions at around 300 ms (e.g., Lewis et al., 2011; Simon et al., 2012). Little is known, however, about the time course of conceptual relations in the brain. Examination of the millisecond-by-millisecond effects of thematic and taxonomic relations on ROI activation could therefore contribute to our understanding of the temporal and spatial nature of the mechanisms of conceptual relations in visual word recognition.

In sum, our study focuses on the role of the ATL and TPJ during processing of taxonomically vs. thematically related word pairs. If these regions represent distinct systems of semantic knowledge, we should find differential neural activation to taxonomic vs. thematic stimuli. Lastly, we examined taxonomic and thematic effects on posterior occipital regions, which Kalénine et al. (2009) found to be involved in taxonomic processing. These regions may be less involved in identifying taxonomic relations with word stimuli.

### 2. Method

#### 2.1. Participants

The MEG experiment included 17 right-handed native English speakers (8 males) from the New York University community with normal or corrected to normal vision. Of the 17 participants, 13 had structural MRI data sets available, which we later used for source localization.

#### 2.2. Stimuli

The authors created the stimulus set via collaboration and agreed upon definitions of each kind of relationship. Taxonomically related items shared a common superordinate category (e.g., cat–wolf) while not having a thematic relation (e.g., items like cat–dog were not used). Thematically related items shared a spatial or functional relationship (e.g., respectively, candles–cake, key–lock). In constructing the thematic pairs, we took care to avoid items sharing salient taxonomic categories, e.g., mouse–cat. The test stimuli included 300 primes and 150 targets. Half of the primes related thematically to the targets (e.g., king–castle), while the other half related taxonomically to the targets (e.g., cottage–castle). The target stimuli could be loosely termed as

---

footnote continued

We suspect the latter source of negative responses can seldom be used in the present task. If we had tested only taxonomic relations, it would be possible to
belonging to one of the following six groups (25 in each): animate/natural objects, clothing, food, tools/artifacts, household, and transportation. We also generated 300 primes and 150 unrelated target filler items (e.g., nutmeg–reflex). Appendix A lists the test stimuli. Association strengths (Nelson et al., 2004) were extremely low for both the taxonomic and thematic pairs (although some association in thematic items is to be expected). None of the filler pairs occurred together in the association norms. Appendix B describes the association strengths of the stimuli in greater detail.

### 2.3. Variables

On average, the words were similar across conditions in terms of length, surface (written word) frequency, bigram (adjacent letter) frequency, number of syllables, number of morphemes, number of phonemes, as well as normative behavioral data including mean naming accuracy as reported by the English Lexicon Project (Balota et al., 2007). Table 1 reports the linguistic properties of the stimuli.

We carefully matched the stimuli in terms of these properties to ensure any effects could be attributed to the category condition rather than lexical properties of the words. Later, we reRESSED these and other properties onto the response data to further ensure the validity of any effects. However, it is important to keep in mind that the same test targets served in the thematic and taxonomic conditions, so any differences between them cannot be explained by lexical effects.

### 2.4. Relatedness ratings

In a separate experiment, we obtained relatedness ratings for the stimulus pairs from 38 native English speakers via Mechanical Turk. We excluded 9 participants based on “sanity-check” items (10 highly related and 10 highly unrelated item pairs, e.g., square–circle and scheme–moose) separate from the experimental stimuli to test whether participants rated the items seriously. We excluded participants when the difference between ratings of highly related and highly unrelated items fell below a criterion value. Instructions for the relatedness task are included in Appendix B. Results of the questionnaire revealed, on average, a higher relatedness score for thematic pairs ($M = 5.63$, $SD = 0.69$) than for taxonomic pairs ($M = 4.47$, $SD = 0.89$). However, both average scores well exceeded that of the filler pairs ($M = 1.82$, $SD = 0.42$). Rather than select a subset of stimuli with equal ratings in the two related conditions, we used each item’s relatedness as a predictor of MEG signal and RT in single-trial analyses.

In the norming study described above, participants were instructed merely to rate the relatedness of the word pairs. We did not provide examples of taxonomic and thematic relation, which were provided in the MEG task (described in the next section). A reviewer raised the concern those examples may have led participants in the main study to have a different understanding of “related.” We therefore obtained a separate set of relatedness ratings using Mechanical Turk in a new task that included the examples of taxonomic and thematic relations used in the MEG task. The task was otherwise identical to the first. We excluded 8 out of 41 participants based on the same criteria for a total of 33 new raters. The old and new ratings were highly correlated ($r = .97$, $p < .001$). Furthermore, the mean ratings for the item types were very similar across the two sets of raters: 4.5 and 4.4 (taxonomic), 5.6 and 5.6 (thematic), and 1.8 and 1.5 (filler). In sum, including examples of different kinds of relations in the instructions did not seem to affect participants’ ratings of the items.

### 2.5. Procedure and recording

We used Matlab (MathWorks, Inc., Natick, MA, USA) with Psychtoolbox helper scripts to present the stimuli. Participants viewed the prime–target pairs over the course of six blocks, with each block consisting of 100 trials. Each block contained 50 filler pairs, 25 taxonomic pairs, and 25 thematic pairs, with randomized presentation. We shifted the block order for each participant and ordered the blocks so that the same prime did not appear twice within the first three sequential blocks. Participants completed a practice session before the actual experiment. They read that they would see word pairs in sequence and should decide whether the pairs were related or unrelated. Instructions explicitly mentioned that related words might be the same kinds of things (like velcro and zipper) or have some relation (like pants and zipper–zippers open pants). They responded yes or no with their left index and middle fingers and were asked to be as fast and accurate as possible. The task lasted approximately half an hour. By requiring participants to use their left hands, we minimized the possibility of movement artifacts in our LH ROIs (which were not motor areas in any case). Furthermore, the analyses focused on activation occurring well before button press. Since both taxonomic and thematic trials used identical positive responses, differences between them cannot be due to response effects.

![Fig. 1. Trial sequence and event durations.](image-url)
Fig. 1 shows the task sequence. The prime word appeared for 300 ms, followed by a variable interval of 0.5–1 s. The target then appeared for 400 ms, so that word exposure did not vary across conditions. The next trial began after the response. Participants lay supine during the experiment while a whole-head MEG system (157 axial gradiometer sensors) (Kanazawa Institute of Technology, Kanazawa, Japan) acquired their neural data. Recording parameters were the following: 1000 Hz sampling rate, 60 Hz bandpass filter, and DC high-pass filter. Structural MRIs of 13 of the participants were acquired from a separate session at the Center for Brain Imaging at New York University (3T Siemens Allegra scanner with T1-weighted MPRAGE sequences).

2.6. Analysis

The source space analysis closely resembled that described in Lewis et al. (2011) and Lewis and Poeppel (2014). First, we noise-reduced the data in MEG160 (Yokogawa Electric and Eagle Technology Corporation, Tokyo, Japan) using reference sensor data and the Continuously Adjusted Least-Squares Method (CALM; Adachi et al., 2001). We next imported the data into MNE (MGH/HMS/MIT Martinos Center for Biomedical Imaging, Charleston, MA) and reconstructed the MRI data sets with FreeSurfer (CorTechs Lab Inc., La Jolla, CA), which we used to calculate minimum-norm solutions. For the four participants without structural MRIs, we employed a standard FreeSurfer brain aligned to each participant’s fiducial data. We calculated estimates of the magnetic field at each sensor (the forward solution) from 5124 points of activity using the boundary-element model (BEM) method. We used the forward solution to estimate the spatio-temporal distribution of the MEG data (the inverse solution). We used a free orientation (unconstrained in relation to the cortical surface) to compute the inverse solution. Data were signed, where negative values indicate activity directed downward and positive values indicate activity directed upward with respect to the head based coordinate space. Finally, data were transformed into dynamic statistical parameter map (dSPM) values (Dale et al., 2000) and retained only components normal to the cortical surface. These analytic procedures (in particular, the use of signed, free orientation minimum norm estimates) follow those used in previous MEG studies of visual and spoken word recognition (e.g., Ettinger et al., 2014; Fruchter et al., 2013; Lewis and Poeppel, 2014; Lewis et al., 2011; Simon et al., 2012). We also applied a standard exclusion procedure to the MEG data to remove movement artifacts (as in previous MEG studies, e.g., Lewis et al., 2011; Simon et al., 2012). For each subject and trial, we summed the number of data points that were two standard deviations above or below the overall mean. We then excluded trials with extreme value counts greater than three standard deviations of the overall extreme mean.

2.7. Regions of interest

Findings from Schwartz et al. (2011) and Kalénine et al. (2009) motivated our interest in four brain regions potentially involved in taxonomic and thematic conceptual processes. For the ATL ROI, this included Brodmann area (BA) regions reported in Schwartz et al. as containing the most significant percentage of voxels associated with taxonomic errors. We therefore created the ATL ROI by merging the following anatomical ROIs from a BA parcellation: BA 18 (33% of error in Schwartz et al.), anterior BA 21 (27% of error), and anterior BA 20 (25% of error). For the TPJ ROI, we merged BA ROIs in posterior regions accounting for the greatest number of voxels associated with thematic errors in Schwartz et al. These included LH BA 39 (35%), posterior LH BA 22 (15%), and LH BA 41 and LH BA 42 (combined 10%). Lastly, we included two additional ROIs encompassing the cuneus and lingual gyrus (LH BA 18 and RH BA 18), reported in Kalénine et al. (2009) as particularly activated in taxonomic categorizations. The ROIs, which were from a parcellation of the standard FreeSurfer brain, were transformed into each participant’s neuroanatomical space. The average activation within the ATL ROI trended negative between 0 and 370 ms post target onset (peaking at 160 ms and 290 ms) and then trended positive until 540 ms (peaking at 430 ms). Average TPJ activation was positive from 50 to 100 ms (peaking at 90 ms) and then remained negative until around 630 ms (peaking at 150 ms and 390 ms). Average LH BA 18 activation was positive from 100 to 700 ms (peaking at 120 ms, 370 ms, and 410 ms). Average RH BA 18 activation was positive from 100 to 510 ms (peaking at 140 ms, 320 ms, and 360 ms). The brain regions and their average time-courses of activation are depicted in Fig. 2.

2.8. Analysis techniques

2.8.1. Behavioral exclusions

We excluded trials with incorrect responses or RTs exceeding 5 s. This removed approximately 9% of the data. After determining average response times in each condition, we converted each participant’s RT data into z-scores and excluded trials in excess of three standard deviations above or below a given participant’s mean. This procedure removed an additional 2% of the data.

2.8.2. Neural analyses

We used two main approaches in the neural analyses: permutation t-tests and single-trial correlations. The former provides a rough estimate of the differences in activation between two conditions (e.g., activation for thematic vs. taxonomic trials). Single-trial correlations provide better evidence of taxonomic vs. thematic processes because they are not a global difference but are graded to the strength of each item’s relation and remove variability due to lexical factors. Ideally, the results of both approaches will agree, but we take as meaningful positive correlations in the absence of the global difference, as the correlation is more sensitive to relation strength (which varied within the groups).

2.8.2.1. Permutation t-test analyses

We conducted two-tailed permutation t-tests to compare differences in ROI activation of taxonomic vs. thematic trials, thematic vs. filler trials, and taxonomic vs. filler trials. We performed the analysis on source space activation over the 200–700 ms window, relative to the target word onset. We did not have strong expectations as to the exact timing of our results. Our time window was chosen to encompass a range across when word meaning becomes available all the way until the end of an analysis epoch. In each test, we set criteria for selecting significant clusters of activation (10 sequential time points with a component test threshold p-value of p = .01). For each cluster of activation, a test statistic was constructed from the sum of t-values within a given cluster to identify the largest cluster statistic. Next, the procedure repartitioned each participant’s data via random assignment to different conditions, and identified the largest cluster statistic at each of the 10,000 permutations, resulting in a distribution of the largest cluster statistics. Comparisons between the original data and the largest cluster statistics created estimates of each ROI’s p-value, corresponding to the average proportion of permutation statistics greater than the original statistic.

Finally, we applied a false discovery rate (FDR) controlling technique (Benjamini and Hochberg, 1995; Benjamini and...
Yekutieli, 2001; Genovese et al. 2002). This involved ordering the p-values in descending order for comparison with the ratio of the index of the ordered p-value to the total number of tests, multiplied by the FDR (here, .05). The p-values less than this statistic were deemed significant.

2.8.2.2. Single-trial analyses. The permutation t-test analyses, which we performed on averaged data, provide a general indication of taxonomic vs. thematic processing differences. We supplemented our analyses with single-trial time-course correlations. This analysis technique, which correlates a given stimulus variable with millisecond-by-millisecond ROI activation across all trials and participants, provides a more in-depth examination of the neural responses implicated in conceptual relations. We primarily focused on neural responses to the following dummy coded variables: Taxonomic vs. thematic (Tax–Them), taxonomic vs. filler (Tax–Fill), and thematic vs. filler (Them–Fill). If a brain region is involved in computing a particular kind of relation, its activation might correlate with the strength of a pair’s relatedness. We therefore included relatedness in each analysis. First, we analyzed ROI activation in regressions to remove any effects of length, bigram frequency, surface frequency, number of morphemes, number of syllables, presentation number (whether first or second time viewing the target), and response time. We next applied a multiple-comparisons correction procedure (Maris and Oostenveld, 2007) to clusters of activation significant at the p < .05 level before correction. Next, we computed an Σr statistic from the summed coefficients of significant contiguous effects and then tested the significance of this statistic with Monte–Carlo p-values. To do so, we created 10,000 permutations of the random variable to generate a correlation wave. Next, we calculated the Σr statistic at each permutation and constructed a new distribution of Σr values from the highest Σr value at each permutation. Finally, we based our Monte–Carlo p-value on the proportion of values greater than the original statistic.

Fig. 2. The ROIs presented on the standard brain (top) and the standard inflated brain (center) presented in lateral, ventral, and occipital views (from left to right). We applied a baseline correction of 100 ms pre-stimulus onset to the average time courses of activation within the ROIs (bottom).
3. Results

The analyses examined behavioral and neural responses to the relatedness scores of each condition and to the dummy coded variables Tax–Fill, Tax–Them, and Them–Fill. First, we report results of our behavioral analysis, which correlated RTs with the variables of interest. We then present results of our permutation t-tests analysis, which provided estimates of the average differences in activation between conditions in each ROI. Finally, we report results of our behavioral analysis, which correlated RTs with the activation within the significant cluster and the error bars show SEMs.

3.1. Behavioral results

On average, thematic targets yielded the fastest responses \( (M=918, SD=497) \), filler targets (with “unrelated” responses) were slowest \( (M=952, SD=528) \), and taxonomic RTs fell in between \( (M=939, SD=517) \). Correlations with RT showed significantly faster responses to taxonomic targets than fillers \( (r=-.03, p=.003) \), significantly faster responses to thematic than fillers \( (r=-.11, p<.0001) \), and significantly faster responses to thematic than taxonomic trials, even though these RTs were measured on the identical target words \( (r=.08, p<.0001) \). Analysis of the RTs from positive trials (thematic and taxonomic) showed a significant correlation with relatedness score, with more related items yielding faster responses \( (r=-.21, p<.0001) \). As expected, RTs from negative trials correlated in the opposite direction, with higher relatedness yielding slower RTs \( (r=.05, p=.0017) \).

Participants responded least accurately to taxonomic items (83% correct), most accurately to fillers (95% correct), and somewhere in between to thematic items (88%). These results seem to reflect a slight bias to respond “unrelated.”

The faster RTs of thematic over taxonomic trials are to be expected, given that the former were more highly related. A regression including relatedness score of each pair along with its condition determined there was no longer a significant difference between conditions when this difference was controlled for. Thus, taxonomic judgments were not slower per se.

Processing speed of taxonomic and thematic judgments may differ depending on whether the object is an artifact like hammer or a natural kind like cherry. Kalénine et al. (2009) found speeded thematic judgments for artifacts and speeded taxonomic judgments for natural kinds. We therefore coded the data into natural kinds and artifacts (counting items from the animate and food groups as natural) and ran separate correlations with the taxonomic and thematic RTs. Like Kalénine et al. (2009), we found that taxonomic responses were significantly faster for natural items than artifacts \( (r=-.05, p<.02) \). This effect was stronger after regressing relatedness score onto RT \( (r=-.08, p<.0001) \). Thematic RT did not reveal any significant correlations.

The main goal of this study was not to compare the speed of the two related conditions but rather to investigate their computation in the brain. In the next section we address that issue.

3.2. Neural results

3.2.1. Neural results: permutation t-tests

We began by analyzing averaged data in each ROI with two-tailed permutation t-tests to determine whether contiguous clusters of activation differed significantly between conditions. The analyses compared activation for taxonomic vs. thematic, taxonomic vs. filler, and thematic vs. filler trials. We report the activation means in dSPM units (Dale et al., 2000). Fig. 3 shows significant results of the permutation t-tests.

3.2.1.1. Anterior temporal. The two-tailed permutation t-tests identified a significant cluster of ATL activity \( (296–614 \text{ ms}, p=.0079) \) for taxonomic vs. filler, with stronger amplitude for filler items than taxonomic items \( (-.671 \text{ filler mean } [.192 \text{ sd}] vs. -.223 \text{ taxonomic mean } [.125 \text{ sd}]) \).

\footnote{We did not find significant effects of natural kinds vs. artifacts in the neural analyses for any of the conditions, perhaps because of insufficient numbers of trials to reveal MEG effects.}
The test did not identify significant activation and Tax
3.2.1.1. Anterior temporal. postime points, the
The Tax
related items likely reflect more processing, as supported by the
behavioral results, where filler items yielded the slowest RTs.

### 3.2.1.2. Temporoparietal.

**Variable** & **ATL** & **TPJ** & **LH BA 18** & **RH BA 18** & **Range** & **Range** & **Range** & **Range**
---
| Them–Fill | 0.09 | 597–600 | 10.71*** | 356–599 | 1.15 | 452–459 | 0.37 | 472–485 |
| Tax–Them | 1.89* | 308–356 | 0.84 | 263–284 | 0.71 | 235–256 | 0.36 | 235–271 |
| Taxonomic relatedness | 4.17** | 419–502 | 1.18 | 481–503 | – | – | 2.94* | 446–497 |
| Thematic relatedness | – | – | 2.56* | 308–355 | 1.90 | 265–302 | – | – |
| Filler relatedness | 0.69 | 200–218 | 4.16** | 365–472 | – | – | 2.21* | 254–310 |

*Note. The r-values are sums across significant clusters. The table shows values and clusters (in ms) that were significant following CMC.

*  p < .05.
**  p < .01.
***  p < .001.

**Table 2**

<table>
<thead>
<tr>
<th>Variable</th>
<th>ATL</th>
<th>TPJ</th>
<th>LH BA 18</th>
<th>RH BA 18</th>
<th>Range</th>
<th>Range</th>
<th>Range</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Them–Fill</td>
<td>0.09</td>
<td>597–600</td>
<td>10.71***</td>
<td>356–599</td>
<td>1.15</td>
<td>452–459</td>
<td>0.37</td>
<td>472–485</td>
</tr>
<tr>
<td>Tax–Them</td>
<td>1.89*</td>
<td>308–356</td>
<td>0.84</td>
<td>263–284</td>
<td>0.71</td>
<td>235–256</td>
<td>0.36</td>
<td>235–271</td>
</tr>
<tr>
<td>Taxonomic relatedness</td>
<td>4.17**</td>
<td>419–502</td>
<td>1.18</td>
<td>481–503</td>
<td>–</td>
<td>–</td>
<td>2.94*</td>
<td>446–497</td>
</tr>
<tr>
<td>Thematic relatedness</td>
<td>–</td>
<td>–</td>
<td>2.56*</td>
<td>308–355</td>
<td>1.90</td>
<td>265–302</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Filler relatedness</td>
<td>0.69</td>
<td>200–218</td>
<td>4.16**</td>
<td>365–472</td>
<td>–</td>
<td>–</td>
<td>2.21*</td>
<td>254–310</td>
</tr>
</tbody>
</table>

This experiment focused on the spatio-temporal nature of neural responses implicated in different kinds of conceptual relations. Our main goal was to test hypotheses from Schwartz et al. (2011) regarding the involvement of the ATL and TPJ in taxonomic and thematic processing. We additionally tested the involvement of posterior occipital regions (LH and RH BA 18) in conceptual processing based on Kähnle et al.’s (2009) findings of occipital involvement in taxonomic classification with picture stimuli. In general, our experiment provides converging evidence of the ATL’s involvement in processing taxonomic conceptual relations, as found by Schwartz et al. (2011). Our results suggest involvement of the TPJ in both relations. The posterior occipital areas showed some evidence of involvement in taxonomic processing. As mentioned in the Method, the permutation t-tests provide a rougher estimate of the differences in conceptual processing than do the single-trial correlations. The permutation t-test has the appropriate sensitivity for comparing the differences between conditions (e.g., early visual responses to long versus short words), but we believe that the correlational analysis was more sensitive to our particular manipulation of continuously varying relatedness.
4.1. Taxonomic specificity in the ATL

Our results imply a distinct role of the ATL in conceptual processing, in agreement with Schwartz et al. (2011). We found modulation of ATL activation by the different conditions beginning as early as 300 ms, which aligns nicely with the timing of word recognition effects in temporal regions found in previous MEG visual word recognition studies (e.g., Lewis et al., 2011; Simon et al., 2012).

First, Tax–Fill modulated the ATL response, with unrelated items resulting in more activation. The direction of this effect presumably reflects greater effort spent in attempting to identify a relation among unrelated stimuli, related to the longer RTs for unrelated responses. In contrast, when items are very similar, identification of relatedness requires comparatively little information retrieval and comparison. We also found a distinction for Tax–Them, with taxonomic items yielding stronger ATL activity. While stronger activation signifies greater processing difficulty, it also corresponds with the number of activated features (Lewis and Poeppel, 2014). Another way to put this is that thematic information largely did not involve the ATL: We did not find effects of thematic vs. filler, thematic relatedness, or filler relatedness on ATL activation in the single-trial analysis. Thus, the Tax–Them difference here probably reflects the lack of thematic computations in the ATL. Permutation t-tests identified a significant difference in activation for taxonomic vs. filler trials, with stronger activation for unrelated items (compatible with results of the single-trial analysis). The test also revealed a significant difference for thematic vs. filler with, again, stronger activation for filler trials. As mentioned earlier, the permutation t-tests give rough estimates of differences between conditions, while the single-trial correlations are more appropriate for measuring effects of continuously varying relatedness on neural responses. Although we cannot rule out the possibility of the ATL’s involvement in thematic relations, our results overwhelmingly emphasize its role in taxonomic processes.

4.2. Thematic processes in the TPJ?

Our results suggest a less specific role of the TPJ in conceptual processing. While we did not find significant differences between conditions in the permutation t-test analysis of the TPJ (other than a suggestive difference for thematic vs. filler), we did find numerous effects in the single-trial correlational analysis not specific to thematic processing. We found correlations between TPJ activation and Them–Fill as well and with Tax–Fill during roughly the same time windows (in both correlations, unrelated items generated stronger activation). Unlike the ATL, the TPJ did not distinguish between taxonomic vs. thematic items. We additionally found that the TPJ correlated with both thematic and filler relatedness, but not with taxonomic relatedness during any time window. Thus, thematic processing seems to have an edge over taxonomic processing here, but the latter cannot be said to be absent.

Although the finding of taxonomic activation (the Tax–Fill effect) in TPJ seems to conflict with Schwartz et al.’s findings, it is not surprising that this region is involved in conceptual representation and processing. In their meta-analysis of 120 imaging studies, Binder et al. (2009) identified the angular gyrus (roughly BA 39, contained in our TPJ) as the area most often found to be activated.
in semantic processing tasks. (They also identified the lateral and ventral temporal cortex as a popular area in such tasks.) Schwartz et al. correlated brain areas with residual errors of the two types after taking into account shared errors; this possibly made it difficult to detect any regions that consistently caused both types of error. Indeed, their TPJ patients made both taxonomic and thematic substitutions, as all but two of their patients made more taxonomic than thematic errors. We address naming errors in general in Section 4.4.

4.3. Occipital processing

We did not find differences between conditions in the permutation t-test analyses involving the posterior occipital ROIs. The correlational analyses, however, revealed an effect of Tax–Fill on the LH ROI, and effects of taxonomic and filler relatedness on the RH ROI. Relative to Kalénine et al.’s (2009) results, our taxonomic effects on occipital responses were far less robust. The occipital effects were also numerically smaller (see Table 2) and not as long lasting as those in the ATL and TPJ. Unlike Kalénine et al., we did not find that occipital activation was significantly stronger for natural taxonomic items, possibly because of the disproportionate number of natural to object/artifact items. The fact that our occipital effects all involved taxonomic (and filler) relatedness is consistent with the idea that visual features used in taxonomic computations are stored in the occipital cortex. Fruit tend to have stems and rounded shapes; animals tend to have legs and heads. These visual commonalities help one to identify categories as taxonomically related. In contrast, thematic relations typically do not involve similar perceptual properties. However, to the degree that one can compare fMRI and MEG results, it appears that Kalénine et al.’s results were rather stronger in the occipital lobe. This is likely because their task involved comparison of pictures. It would be interesting to see whether their taxonomic occipital effects would be found with auditorily presented category names.

4.4. Relation to naming errors

Throughout we have compared our results to those of Schwartz et al.’s study of naming errors in aphasia. In the Introduction, we raised the issue of how lesions in an area relate to errors. According to a simplistic analysis, if a brain region does X, then one would expect to find people with lesions in that region doing less of X. However, Schwartz et al. found that damage in the ATL led to more taxonomic errors, whereas damage in the TPJ led to more thematic errors, concluding that they are involved in taxonomic and thematic processing, respectively.

The argument for the ATL is fairly straightforward. As Schwartz et al. (2011, p. 8522) propose (and also see Patterson et al., 2007), when the ATL is damaged, information about category features is lost, causing errors in labeling: “...the left ATL communicates feature and category information to posterior lexical-phonological systems, a specialization it derives from being part of the bilateral anteromedial and inferolateral network for visual object identification.” If someone loses the information that poodles have curly fur, then he or she might well label a poodle with the name of a straight-haired dog, such as collie, which matches the picture in the preserved features. Such an effect is confirmed in simulations of semantic memory (Rogers and McClelland, 2004). Schwartz et al. (2011) also suggest that the loss of features may be
directional, in that activating a lemma (in naming) is more disrupted than understanding a word, because they controlled for semantic comprehension in their analyses and yet ATL damage predicted taxonomic errors. However, their multiple-choice comprehension test may have been easier than the naming task, so this issue requires further investigation.

It is not as easy to explain thematic errors, however. If the TPJ contains information about thematic relations (as our data suggest), and if damage to that region causes noise in the system, as Schwartz et al. argue, why would aphasic patients provide thematically related names? Surely if their ATL is intact, they would give the correct name; if it is slightly damaged, it would be more correct to give a taxonomic response. There is a sense in which thematic responses are clearly not the right kind of answer – "bone" is not an appropriate name for a picture of a dog. One potential answer is that the TPJ is involved in regulating language use, in addition to its role in conceptual processing. Noonan et al. (2009) argue that this region is part of a language regulation system (including prefrontal cortex) that coordinates sentence processing and semantic activation. In particular, when reading ambiguous words, it would help to assess thematic information to aid disambiguation (e.g., money-bank vs. fishing-bank), but speakers do not want to produce thematic information instead of the target word, so such associates must be inhibited in production. Noonan et al. (2009) propose that a number of aphasic symptoms following damage to the TPJ could be caused by failure to inhibit related semantic information.

Although this notion seems useful in explaining thematic naming errors, our own results suggest thematic knowledge is also processed in the TPJ, given that our task did not involve production or sentence comprehension but merely judgments of relatedness. Semantic control does not seem to be a critical part of that task – participants were merely judging whether two items were related. They did not have to select or filter out any relation. Therefore, we propose that the TPJ is involved in detecting thematic relations and also in some form of attentional regulation, perhaps related to controlling the use of those relations in language processing. Loss of control results in thematically related names, especially when the correct name is not immediately forthcoming.

That said, the results across all the studies are complex, and no single answer can yet explain all the data involving the TPJ. Schwartz et al. (2011) argued that loss of inhibitory control would predict greater numbers of "off-task" responses; instead, such responses were associated with ATL damage in their study. However, Schwartz et al. describe most such responses as phrases that described rather than named the target, for example, "it goes neigh" for the picture of the horse. It is possible that such responses merely reflect the loss of the name after ATL damage, with the participant attempting to identify the object even when the name is not available. Of course, this idea is speculative at this point, but it highlights the fact that more data are needed regarding the processing in this area, including a comparison of different tasks.

The demands of picture naming are not the same as those of judging whether words are related, so some differences in the tasks’ neural processing is to be expected.

4.5. Summary

Our results support theories of the ATL’s importance in representing semantic knowledge (e.g., Rogers et al., 2004; Tyler and Moss, 2001). Our finding of differential modulation of the ATL and TPJ for taxonomic vs. thematic items in general corroborates Schwartz et al.’s (2011) proposal for two distinct systems for representing different kinds of semantic knowledge. Our results provide compelling evidence of the ATL’s involvement in representing taxonomic knowledge, and only weak evidence of its sensitivity to thematic concepts. Conversely, we found strong evidence of thematic processing in the TPJ. Computations in both regions occurred during similar time windows (beginning as early as 300 ms). The specificity of the TPJ for thematic knowledge, however, is less clear, as it responded to Them–Fill and thematic relatedness, to Tax–Fill, and also filler relatedness. Also, unlike for the ATL, we did not find differences in TPJ activation between conditions. This pattern of results suggests that the ATL plays a strong role in taxonomic judgments, while the TPJ’s role appears more generic.

In the Introduction, we raised the question as to whether taxonomic and thematic information are represented together. The answer seems to be both yes and no. ATL activation seemed closely predicted by taxonomic relations, with no influence of thematic relations in the single-trial analyses. The importance of the ATL to taxonomic knowledge is also clear from studies of semantic dementia (e.g., Lambon Ralph et al., 2001; Mesulam et al., 2009). However, the TPJ showed influences of both kinds of relations. As we remarked earlier, thematic relations may be based on events and situations, but the knowledge of those events is about taxonomic categories: what kinds of entities are present and how the entities interact with one another. Therefore there must be some representation that links co-occurring taxonomic categories. Understanding the thematic relation of dogs to bones requires one to represent the taxonomic categories of dogs and bones. Consistent with Schwartz et al.’s (2011) findings, those connections seem to be made in the TPJ.

Acknowledgments

This work was supported by National Science Foundation grants DGE-1342536 to GL, BCS-1128769 to GM, and by National Institute of Health grant 2R01DC05660 to DP. The authors thank Alec Marantz and Joseph Fruchter for providing helpful advice, Patricia Chan and Anna Coenen for technical help, and Laura Gwilling for help with the permutation analyses.

Appendix A. Test stimuli (taxonomic prime/thematic prime → target)

<table>
<thead>
<tr>
<th>Animate</th>
<th>Food</th>
<th>Dishes/appliances</th>
</tr>
</thead>
<tbody>
<tr>
<td>psychic/play</td>
<td>mango/peeler → apple</td>
<td>bbell/fire → alarm</td>
</tr>
<tr>
<td>teacher/mask</td>
<td>cracker/bakery → bread</td>
<td>toaster/smoothie → blender</td>
</tr>
<tr>
<td>seal/dam</td>
<td>pork/ketchup → burger</td>
<td>squeegee/debris → broom</td>
</tr>
<tr>
<td>gymnast/gloves</td>
<td>yogurt/bun → butter</td>
<td>vial/water → canteen</td>
</tr>
<tr>
<td>centipede/net</td>
<td>scone/candies → cake</td>
<td>furnace/smoke → chimney</td>
</tr>
<tr>
<td>wolf/lap</td>
<td>muffin/bowl → cereal</td>
<td>spatula/pasta → colander</td>
</tr>
</tbody>
</table>
Nelson et al. (2004). The corpus listed values for 101 of the taxonomic, 123 of the thematic, and 137 of our Appendix B. Association strengths of the stimuli

crouton/dip → chips
raisin/box → chocolate
brandy/thermos → cocoa
grape/palm → coconut
nectar/mug → coffee
ale/bottle → cola
candy/wrapper → gum
liquor/straw → juice
soda/pitcher → lemonade
clam/trap → lobster
pea/tarars → onion
tomato/brine → pickle
casserole/plate → sandwich
herring/can → sardines
porridge/spoon → soup
rice/fork → spaghetti
ham/grill → steak
dumpling/soy sauce → sushi
vodka/sugar → tea
speakers/program → computer
pan/stew → crock
rag/chalk → eraser
teapot/whiskey → flask
pot/coaster → glass
rod/bait → hook
bin/jam → jar
canister/beer → keg
fryer/whistle → kettle
flashlight/book → lamp
latch/key → lock
recipe/diner → menu
wood/rust → metal
microwave/food → refrigerator
carafe/cup → saucer
blemish/bleach → stain
jug/ashes → urn
goblet/bouquet → vase
toilet/bath → tub

Clothing

bib/chef → apron
medal/cop → badge
mascara/cheek → blush
vein/fracture → bone
vault/clothing → closet
barrette/hair → comb
broach/gown → corsage
cot/infant → cradle
camera/eyes → glasses
dishes/hamper → laundry
detergent/skin → lotion
knot/neck → noose
gold/oyster → pearl
baton/cheerleader → pompon
sombrero/storm → poncho
string/braid → ribbon
bracelet/finger → ring
slip/throat → scarf
blouse/lady → skirt
robe/artist → smock
skates/laces → sneakers
mittens/legs → stockings
uniform/groom → tuxedo
rainboots/rain → umbrella
turban/bride → veil

Tools/miscellaneous

crutch/wound → bandage
clipboard/leaflet → binder
fog/drizzle → cloud
dollar/mercer → coin
hammock/toddler → crib
pamphlet/padlock → diary
locker/socks → drawer
package/letter → envelope
emblem/pole → flag
yarn/teeth → floss
sword/bullet → gun
drill/nail → hammer
lighter/candle → match
marker/note → pencil
quilt/head → pillow
spade/leaves → rake
scissors/beard → razor
magazine/actor → script
brick/tar → shingle
hoe/earth → shovel
shampoo/cloth → soap
smudge/postcard → stamp
painting/chisel → statue
staple/paper → tape
lane/map → route

Transportation/places

surfboard/baby → stroller
taxi/stewardess → airplace
speedboat/stretcher → ambulance
skateboard/road → bike
buggy/tourist → bus
yacht/oar → canoe
tractor/garage → car
cottage/king → castle
valley/bat → cave
resort/nurse → clinic
stairs/passenger → elevator
crypt/tombstone → grave
truck/pilot → helicopter
alley/sedan → highway
ship/airport → jet
sea/kayak → lake
song/student → lecture
tournament/runner → marathon
scooter/grass → mower
bazaar/lunatic → sanctions
convertible/aeonaut → spaceship
aisle/jeep → street
raft/ocean → submarine
motorcycle/rail → train
aquarium/lion → zoo

Appendix B. Association strengths of the stimuli

We checked the association strengths of our prime–target pairs against the list of South Florida free association normative data from Nelson et al. (2004). The corpus listed values for 101 of the taxonomic, 123 of the thematic, and 137 of our filler primes. Nineteen taxonomic targets, 49 thematic targets, and no filler targets were listed as responses to these primes. The “forward cue to target strength” of the pairs was low for both taxonomic pairs (M = .01, SD = .02) and thematic pairs (M = .06, SD = .13).

Means (SDs) of association strengths by condition

<table>
<thead>
<tr>
<th>Measure</th>
<th>Taxonomic</th>
<th>Thematic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward cue to target strength</td>
<td>.01 (.02)</td>
<td>.06 (.13)</td>
</tr>
<tr>
<td>Backward target to cue strength</td>
<td>.02 (.05)</td>
<td>.15 (.21)</td>
</tr>
<tr>
<td>Mediated strength</td>
<td>.04 (.07)</td>
<td>.01 (.02)</td>
</tr>
<tr>
<td>Overlapping associated strength</td>
<td>.05 (.06)</td>
<td>.03 (.06)</td>
</tr>
</tbody>
</table>
Instructions for Mechanical Turk relatedness task (Study 1)

People can make various predictions about objects, people, or animals after reading a short story about them. Today we would like you to make judgments about objects, people, or animals outside the context of a story. You will be shown a pair of words. We would like you to rate the relatedness of the two words on a scale of 1–7. 1 means the words are “not at all related” and 7 means the words are “highly related.” Make your decision by pressing the corresponding number key on your keyboard. Please use all the numbers in the 1–7 scale, not just one or two. There is no right or wrong answer to these questions—we’re just interested in your opinion. You have as much time as you want, but it is usually best to just go with your first reaction about how related the words are.

Instructions for Mechanical Turk relatedness task (Study 2)

In this experiment, you will be shown pairs of words. We would like you to rate the relatedness of the two words on a scale of 1–7. We consider words to be “related” if they are the same general kind of thing (like velcro and zipper-both are kinds of fasteners) or if they are related to one another (like pants and zipper-zippers open pants). Examples of unrelated items would include words like desk and harp, or cafe and harp. Of course, words can be more or less related, and we would like you to decide just how related (if at all) each pair of words is. 1 means the words are “not at all related” and 7 means the words are “highly related.” Make your decision by pressing the corresponding number key on your keyboard. Please use all the numbers in the 1–7 scale, not just one or two. There is no right or wrong answer to these questions—we’re just interested in your opinion. You have as much time as you want, but it is usually best to just go with your first reaction about how related the words are.

Instructions for MEG related task

In this experiment, you will view pairs of related and unrelated word pairs. We consider words to be “related” if they are the same general kind of thing (like velcro and zipper-both are kinds of fasteners) or if they are related to one another (like pants and zipper-zippers open pants). Examples of unrelated items would include words like desk and harp, or cafe and harp. Each trial will commence with a fixation cross. Look at that cross. Next, a word will appear, followed by a second word. Respond “yes” with your index finger if they are related. Respond “no” with your middle finger if they are not related. Please respond as quickly and as accurately as possible.

References


