

Journal Club

Editor's Note: These short reviews of recent *JNeurosci* articles, written exclusively by students or postdoctoral fellows, summarize the important findings of the paper and provide additional insight and commentary. If the authors of the highlighted article have written a response to the Journal Club, the response can be found by viewing the Journal Club at www.jneurosci.org. For more information on the format, review process, and purpose of Journal Club articles, please see <http://jneurosci.org/content/preparing-manuscript#journalclub>.

Challenges in Studying Multidimensional Semantic Representations in the Human Brain

Colleen Mills-Finnerty

Mental Illness Research, Education, and Clinical Care, Veterans Administration Palo Alto Health Care System, Department of Psychiatry and Behavioral Science, Stanford University, Palo Alto, California 94304-1419

Review of Xu et al.

Concepts are a necessary component of higher-order cognition, and they facilitate the comprehension and production of language. However, because semantic representations are encoded by spatially diverse populations of neurons across frontal, parietal, and temporal lobes, localizing function is difficult (Huth et al., 2016). Adding to the challenge, the similarity of semantic stimuli can be delineated using multiple conceptual dimensions (e.g., shared features, co-occurrence, familiarity), which may be represented by different patterns of distributed neuronal response. These factors have hindered our understanding of how concepts are represented by the brain.

Research on semantic encoding suggests several hypotheses about how conceptual representations are implemented by the brain. One proposal is that concepts are encoded through distributed, anatomically connected, modality-specific regions specialized to process visual, auditory, or spatial features of stimuli (Tyler and Moss, 2001; Wang et al., 2016). Another theory proposes that the neural mechanisms of semantic

processing involve a “hub and spoke” configuration, in which a “hub” region of the brain serves as a general purpose intermediary to many inputs and outputs that respond in a specialized manner to particular features of stimuli (Rice et al., 2015; Ralph et al., 2017). The key difference between these theories is whether a centralized hub region is necessary for generating coherent, integrated representations of semantic concepts, or if distributed, decentralized patterns of activation are sufficient.

The anterior temporal lobe (ATL) has been proposed as a hub integrating inputs from “spoke” regions involved in conceptual processing. Evidence for this includes findings that ATL lesions lead to increased rates of category-independent semantic errors (Schwartz et al., 2011); ATL degeneration leads to semantic dementia which is characterized by word naming errors (Ralph et al., 2017); and intracranial recordings in ATL showed modality-independent coding during proper name retrieval and picture naming (Abel et al., 2015; Shimotake et al., 2015). The ATL may also contain graded representations of more focal types of semantic relations, facilitated by cytoarchitectural gradients and anatomical connectivity between ATL and regions specialized for processing sensory features, such as visual or auditory information (Ding et al., 2009; Binney et al., 2010). For example, focal ATL damage in herpes simplex encephalitis leads to category specific

deficits in naming biological entities, possibly mediated by deficits in white matter connectivity between ATL and visual regions (Gainotti, 2018).

To clarify the role of potential hub regions in semantic categorization, a recent study by Xu et al. (2018) used fMRI to investigate brain responses during categorization focusing on two dimensions of semantic similarity: taxonomic categorization (stimuli with similar features; e.g., people or objects) and thematic (stimuli that co-occur; e.g., people, places and things associated with sports or school). Participants completed a simple task in which they categorized the same set of words along either taxonomic or thematic dimensions. The authors used representational similarity analysis (RSA) to compare the brain states associated with each form of categorization. RSA uses representative geometry to estimate the information contained within voxel patterns of activation as they relate to different task conditions (Kriegeskorte and Kievit, 2013). Since fMRI has limited temporal resolution and summarizes activation indirectly over very large populations of neurons (on the order of thousands to millions), multivariate methods such as RSA can provide an intermediate descriptive level between neuronal populations and cognitive representations. Xu et al. (2018) examined taxonomic and thematic processing using RSA in two regions of

Received May 24, 2018; revised June 28, 2018; accepted June 29, 2018.

I thank Drs. Manjari Narayan and Catherine Hanson for their helpful comments on earlier versions of this paper.

The author declares no competing financial interests.

Correspondence should be addressed to Dr. Colleen Mills-Finnerty, Department of Psychiatry and Behavioral Science, Stanford University, 401 Quarry Road, Palo Alto, CA 94304-1419. E-mail: cmfynn@stanford.edu.

DOI:10.1523/JNEUROSCI.1354-18.2018

Copyright © 2018 the authors 0270-6474/18/387029-03\$15.00/0

interest, the left ATL and the temporoparietal junction (TPJ). They chose these areas based on results from a large study in stroke patients, which implicated lesions of ATL in increased taxonomic naming errors, and TPJ in thematic naming errors (Schwartz et al. 2011). They also conducted whole-brain searchlight analysis of taxonomic and thematic processing, RSA within two a priori defined semantic networks, and multidimensional scaling (MDS) to map the representational semantic space within ATL and TPJ.

Xu et al. (2018) identified a complex pattern of encoding in the ATL and TPJ, with both regions representing thematic and taxonomic information, but to different degrees. Using RSA, they found that both the left ATL and TPJ regions of interest responded more strongly to taxonomic categorization than to thematic, but TPJ showed a greater response to thematic categorization than ATL. They used Spearman's rank partial correlation to measure representational patterns in the brain during taxonomic categorization while controlling for effects of thematic categorization, and vice versa. Using this approach, they found that responses in ATL during thematic categorization were significant only when controlling for taxonomic categorization, whereas TPJ significantly represented both taxonomic and thematic categorization. These findings suggest that populations of neurons in ATL and TPJ represent both thematic and taxonomic relations, but thematic relations are embedded within taxonomic relations. Another interpretation is that patterns of neuronal firing in response to shared features may lead to more stable patterns for the similarity-based taxonomic category compared with the co-occurrence based thematic category (see limitations discussion below).

Whole-brain searchlight analyses revealed a similarly complex pattern, where brain response to thematic processing was stronger when controlling for taxonomic processing, and vice versa. A distributed network of regions across the parietal, frontal, and occipital cortex responded during taxonomic categorization, and this network corresponded well with foci identified in a meta-analysis of semantic processing (Binder et al., 2009). When examining taxonomic categorization and controlling for thematic, these results became stronger and more distributed. Notably, however, ATL did not show a significant response during taxonomic categorization measured using whole-brain searchlight analysis.

Based on lesion results and hub and spoke theories of semantic processing, the ATL has been proposed to function as a hub that integrates inputs from other regions to facilitate behaviors such as categorization. However Xu et al. (2018) only detected significant ATL responses during taxonomic categorization when using RSA at the ROI level, and then only with a modest Spearman's $r = 0.24$. ATL responses were not significant during taxonomic categorization when measured using whole-brain searchlight analysis. These findings suggest that while some neurons in ATL represent taxonomic information, this representation is not as strong as in other regions in the whole-brain analysis. This calls into question whether ATL meets the definition of a hub in this context. Thus, the results seem to better support the distributed theory of semantic network function. Still, it is possible that the hub role of the ATL would be delineated more clearly using metrics intended to capture how essential a region is to the function of the semantic network, such as connectivity analysis (Hoffman et al., 2015; Jung and Lambon Ralph, 2016; Wang et al., 2016), or with imaging sequences better optimized for signal-to-noise in this region (see meta-analysis by Visser et al., 2010). The ATL might show strong connectivity in concert with regions that encode modality-specific stimulus features to facilitate coherent conceptual representations, despite not showing a strong response on its own.

Overall, these findings suggest that in healthy young adults, taxonomic and thematic categorization are not well behaved processes that correspond neatly to distinct neural substrates, despite what lesion studies and theories of semantic encoding might suggest (Schwartz et al., 2011; Ralph et al., 2017). Semantic categorization instead appears to rely on a complex, distributed, multivariate signal where discrete regions represent multiple dimensions of categorization even while preferring certain dimensions over others. The authors interpret their findings of greater representation of taxonomic relations in ATL and TPJ relative to thematic processing to mean that thematic processing is subordinate to, or embedded within, taxonomic processing. However, there are several concerns that limit interpretation of these findings. First, there appears to be variability in the typicality of word stimuli within their taxonomic or thematic domain. For example, the place words associated with sports, such as stadium, award platform, etc. could also apply to school events. This is illustrated by the results of multidimensional scaling, which was used to measure the distance in representational

space between types of stimuli; stimuli that are closer in representational space tend to recruit more similar forms of processing than stimuli that are further apart. Within ATL and TPJ, sports location stimuli grouped more closely to school and medical locations than to sports-related people or object stimuli. Similarly, the school object stimuli such as chalk and ruler can map onto other dimensions (e.g., art supplies), and group more closely to other objects rather than the other school stimuli in the MDS plots, suggesting a lack of specificity of stimulus-category mappings. If participant perception of stimulus-category ambiguity contributed to variability in strength of neuronal representations, accounting for this explicitly might reveal more stable representational patterns, such as stronger thematic categorization in TPJ, without needing to first account for taxonomic processing. This might help clarify whether thematic processing is truly embedded within taxonomic processing in the brain, and whether ATL and TPJ serve as hubs or as part of a decentralized distributed network.

The findings of Xu et al. (2018) also highlight challenges of interpreting differences in neuronal responses while comparing taxonomic and thematic categorization (Mirman et al., 2017). Thematic stimuli are linked to each other based on co-occurrence, and experiences of co-occurrence vary across individuals. This may lead to discrepancies in the stability of stimulus representations across participants and interfere with the reliability of group level inference, as has been demonstrated for representations of other concepts that vary in stability (e.g., concrete versus abstract categorization; Anderson et al., 2014). This potential confound adds ambiguity as to whether there is truly an embedded structure of thematic relations within taxonomic relations in ATL and TPJ. This could be clarified by testing whether within- versus between-subject classification accuracy varies based on taxonomic and thematic categorization. More generally, these results highlight challenges in interpreting brain response during laboratory experiments, which may not generalize to the more complex naturalistic context (Krakauer et al., 2017). There are a wide range of approaches used to study semantic categorization, ranging from the simple, hypothesis-driven experimental approaches such as Xu et al. (2018) to the data-driven analysis of hours of passive listening of Huth et al. (2016). It remains challenging to find a middle ground between experimental control in the laboratory and the ecological validity of capturing a process close to how it happens in the real world.

In sum, Xu et al. (2018) shed light on how distributed populations of neurons can flexibly represent taxonomic and thematic categorization. Their results suggest a more nuanced understanding of semantic categorization than a 1:1 mapping of function to specific brain regions and demonstrate that even potentially specialized areas of the ATL and TPJ represent categorical dimensions outside their preferred domain. More research will be needed to discover the boundaries of this flexible representation and how these functions subserved real-world behaviors.

References

- Abel TJ, Rhone AE, Nourski KV, Kawasaki H, Oya H, Griffiths TD, Howard MA 3rd, Tranel D (2015) Direct physiologic evidence of a heteromodal convergence region for proper naming in human left anterior temporal lobe. *J Neurosci* 35:1513–1520. [CrossRef Medline](#)
- Anderson AJ, Murphy B, Poesio M (2014) Discriminating taxonomic categories and domains in mental simulations of concepts of varying concreteness. *J Cogn Neurosci* 26:658–681. [CrossRef Medline](#)
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19:2767–2796. [CrossRef Medline](#)
- Binney RJ, Embleton KV, Jefferies E, Parker GJ, Lambon Ralph MA (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. *Cereb Cortex* 20:2728–2738. [CrossRef Medline](#)
- Ding SL, Van Hoesen GW, Cassell MD, Poremba A (2009) Parcellation of human temporal polar cortex: a combined analysis of multiple cytoarchitectonic, chemoarchitectonic, and pathological markers. *J Comp Neurol* 514:595–623. [CrossRef Medline](#)
- Gainotti G (2018) Why do herpes simplex encephalitis and semantic dementia show a different pattern of semantic impairment in spite of their main common involvement within the anterior temporal lobes? *Rev Neurosci* 29:303–320. [CrossRef Medline](#)
- Hoffman P, Lambon Ralph MA, Woollams AM (2015) Triangulation of the neurocomputational architecture underpinning reading aloud. *Proc Natl Acad Sci U S A* 112:E3719–E3728. [CrossRef Medline](#)
- Huth AG, de Heer WA, Griffiths TL, Theunissen FE, Gallant JL (2016) Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532:453–458. [CrossRef Medline](#)
- Jung J, Lambon Ralph MA (2016) Mapping the dynamic network interactions underpinning cognition: a cTBS-fMRI study of the flexible adaptive neural system for semantics. *Cereb Cortex* 26:3580–3590. [CrossRef Medline](#)
- Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D (2017) Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93:480–490. [CrossRef Medline](#)
- Kriegeskorte N, Kievit RA (2013) Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn Sci* 17:401–412. [CrossRef Medline](#)
- Mirman D, Landrigan JF, Britt AE (2017) Taxonomic and thematic semantic systems. *Psychol Bull* 143:499–520. [CrossRef Medline](#)
- Ralph MA, Jefferies E, Patterson K, Rogers TT (2017) The neural and computational bases of semantic cognition. *Nat Rev Neurosci* 18:42–55. [CrossRef Medline](#)
- Rice GE, Lambon Ralph MA, Hoffman P (2015) The roles of left versus right anterior temporal lobes in conceptual knowledge: an ALE meta-analysis of 97 functional neuroimaging studies. *Cereb Cortex* 25:4374–4391. [CrossRef Medline](#)
- Schwartz MF, Kimberg DY, Walker GM, Brecher A, Faseyitan OK, Dell GS, Mirman D, Coslett HB (2011) Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proc Natl Acad Sci U S A* 108:8520–8524. [CrossRef Medline](#)
- Shimotake A, Matsumoto R, Ueno T, Kunieda T, Saito S, Hoffman P, Kikuchi T, Fukuyama H, Miyamoto S, Takahashi R, Ikeda A, Lambon Ralph MA (2015) Direct exploration of the role of the ventral anterior temporal lobe in semantic memory: cortical stimulation and local field potential evidence from subdural grid electrodes. *Cereb Cortex* 25:3802–3817. [CrossRef Medline](#)
- Tyler LK, Moss HE (2001) Towards a distributed account of conceptual knowledge. *Trends Cogn Sci* 5:244–252. [CrossRef Medline](#)
- Visser M, Jefferies E, Lambon Ralph MA (2010) Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J Cogn Neurosci* 22:1083–1094. [CrossRef Medline](#)
- Wang X, Fang Y, Cui Z, Xu Y, He Y, Guo Q, Bi Y (2016) Representing object categories by connections: evidence from a multivariate connectivity pattern classification approach. *Hum Brain Mapp* 37:3685–3697. [CrossRef Medline](#)
- Xu Y, Wang X, Wang X, Men W, Gao JH, Bi Y (2018) Doctor, teacher, and stethoscope: neural representation of different types of semantic relations. *J Neurosci* 38:3303–3317. [CrossRef Medline](#)