A comparison of three vector space models of word meaning for mapping the semantic system

VANDERBILT UNIVERSITY®

Deborah F. Levy, Stephen M. Wilson

Department of Hearing and Speech Sciences, Vanderbilt University Medical Center

Introduction

What are the relative utilities of different vector space models of word meaning for mapping the semantic system in the brain? In this preliminary study, we investigated the neural correlates of three distinct models of word meaning – the experiential attributes model (Binder et al., 2016), fastText (Bojanowski et al., 2017) and GloVe (Pennington et al., 2014) – using fMRI and representational similarity analysis (RSA; Kriegeskorte et al., 2008).

Methods

We scanned ten neurologically typical participants with 3 Tesla fMRI as they processed single words in an event-related paradigm. As this was an exploratory study, the mode of presentation (auditory vs. visual), set of words, behavioral task, and stimulus timing varied across participants. All participants completed six runs of their respective tasks.

Vector space representations for each word in each wordlist were constructed using three vector space models: the experiential attributes model, fastText, and GloVe. Representational dissimilarity matrices (RDMs) were then created by calculating the cosine distance between all pairwise feature vectors under each semantic model (Fig. 1).

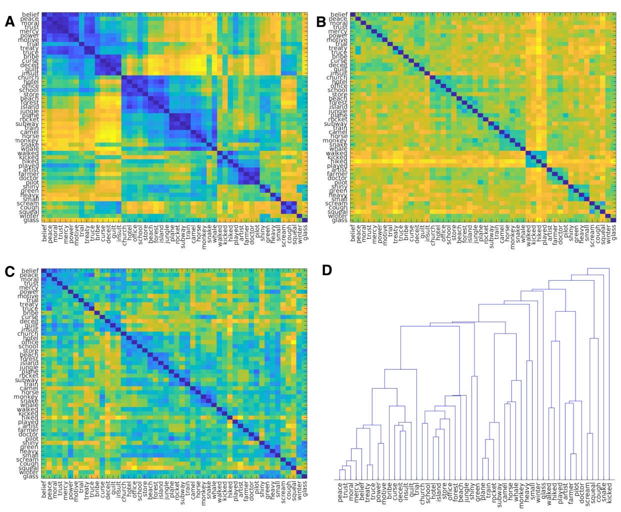

Figure 1. Depictions of semantic distance using different vector space models of word meaning on a wordlist with 50 words. A-C: RDMs for (A) experiential attributes model, (B) fastText, (C) GloVe. D: Dendrogram reflecting hierarchical clustering under the experiential attributes model for the same 50-word wordlist.

A searchlight approach was implemented using an in-house Matlab script to calculate correlations between patterns of semantic similarity and patterns of neural similarity centered at each voxel in the brain. Voxelwise t-tests across participants were performed on the resulting correlation maps to identify regions where model-based semantic similarity and neural similarity were reliably correlated.

Results

Group maps based on each of the three vector space models of word meaning revealed a somewhat left-lateralized semantic network including the bilateral angular gyri and left inferior frontal gyrus, as well as the left middle temporal gyrus for two of the three vector space models (Fig. 2).

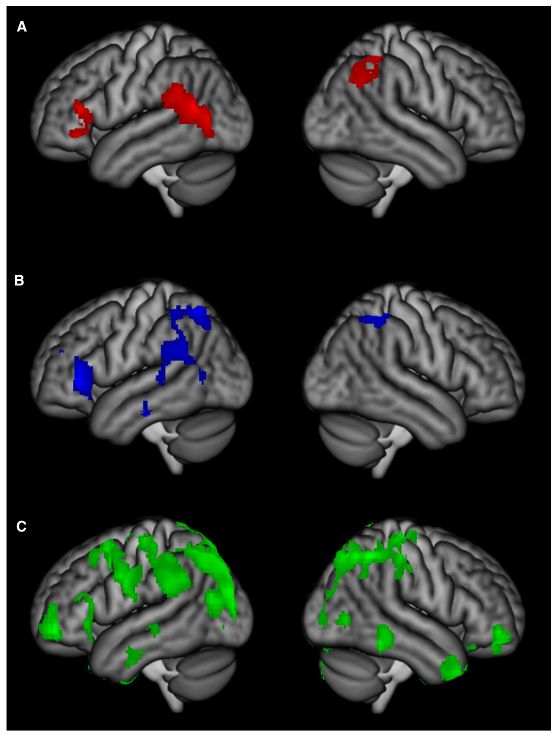

Figure 2. Group maps (n=10) depicting regions where semantic similarity reliably correlated with neural similarity for each vector space model of word meaning. (A) Experiential attributes, (B) fastText, (C) GloVe. Images are thresholded at voxelwise $p<0.01$ and to show only clusters that meet or exceed 3 cm^3 in volume.

A series of paired-sample t-tests revealed no significant differences between maps based on the three different vector space models of word meaning.

Discussion

Semantic maps as generated using all three vector space models of word meaning – the experiential attributes model, fastText, and GloVe – aligned with previous characterizations of the semantic network (Binder et al., 2009).

Prior work has investigated the utility of different vector space models of word meaning in decoding-based analyses in both brains (Xu et al., 2016; Anderson et al., 2016; Abnar et al., 2018) and behavior (Pereira et al., 2016; Baroni et al., 2014). Although no significant differences in the usefulness of varying vector space models of word meaning were detected in our analysis, this may reflect our limited sample size.

Our findings suggest that diverse models of word meaning can be used to identify brain regions that encode semantic representations.

Acknowledgments

This research was supported by the National Institutes of Health (NIDCD R01 DC013270). We thank the members of the Language Neuroscience Lab, the MRI technologists at the Vanderbilt University Institute for Imaging Science, and all of the individuals who participated in the study.

References

1. Binder JR, Conant LL, Humphries CJ, Fernandino L, Simons SB, Aguilar M, et al. Toward a brain-based componential semantic representation. Cognitive Neuropsychology 2016;33:130–74.
2. Bojanowski P, Grave E, Joulin A, Mikolov T. Enriching word vectors with subword information. Transactions of the Association for Computational Linguistics 2017;5:135–46.
3. Pennington J, Socher R, Manning CD. GloVe: Global vectors for word representation. EMNLP, 2014: 1532-1543.
4. Kriegeskorte N, Mur M, Bandettini PA. Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci 2008;2.
5. Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 2009;19:2767–96.
6. Xu H, Murphy B, Fyshe A. BrainBench: A brain-image test suite for distributional semantic models. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Texas: Association for Computational Linguistics; 2016, p. 2017–21.
7. Anderson AJ, Zinszer BD, Raizada RDS. Representational similarity encoding for fMRI: Pattern-based synthesis to predict brain activity using stimulus-model-similarities. NeuroImage 2016;128:44–53.
8. Abnar S, Ahmed R, Mijnheer M, Zuidema W. Experiential, distributional and dependency-based word embeddings have complementary roles in decoding brain activity. Proceedings of the 8th Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2018), Salt Lake City, Utah: Association for Computational Linguistics; 2018, p. 57–66.
9. Pereira F, Gershman S, Ritter S, Botvinick M. A comparative evaluation of off-the-shelf distributed semantic representations for modelling behavioural data. Cognitive Neuropsychology 2016;33:175–90.
10. Baroni M, Dinu G, Kruszewski G. Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Baltimore, Maryland: Association for Computational Linguistics; 2014, p. 238–47.