Unraveling the visual and semantic components of object representation

D.D. Leeds[1], D.A. Seibert[2,3], J.A. Pyles[1] and M.J. Tarr[1,4]

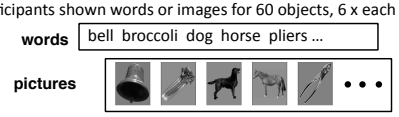
[1]Center for the Neural Basis of Cognition, Carnegie Mellon University (CMU), Pittsburgh, PA, [2]Department of Biomedical Engineering, University of Houston, Houston, TX, [3]uPNC Summer Program, CMU, [4]Department of Psychology, CMU

Components of object "semantics"

- Past studies associate semantic activation with visual *and* non-visual areas, but intermingle picture and word stimuli (*e.g.,* Just et al., 2010)
- BOLD responses were associated with object semantics for pictures vs. words presented in *separate* conditions
- Analyses of neural data included MVPA within a "searchlight" procedure and correlations with stimulus image similarity as measured by a variety of computer vision methods

Methods

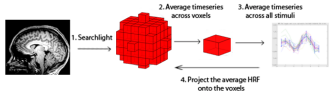
- Participants shown words or images for 60 objects, 6 x each



- BOLD signals recorded with slow event-related design (2 sec TR, partial coverage)

Searchlight analysis

- Constructed "searchlight"—123 voxel sphere—centered at each voxel (Kriegeskorte et al., 2006)

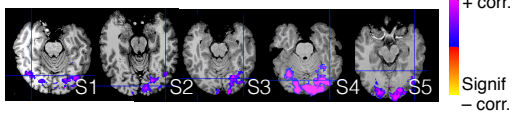


- Compared voxel population responses within and between classes (pictures vs. words)

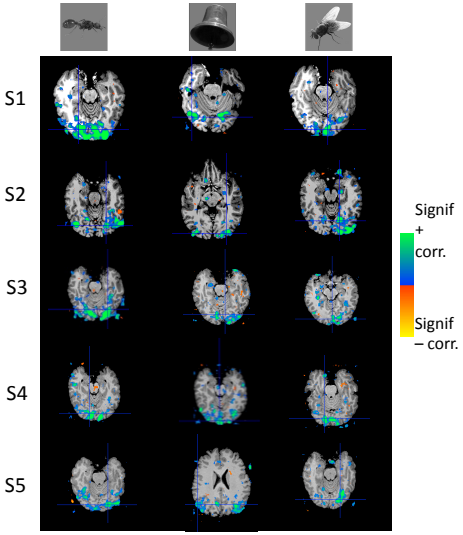
$$\text{contrast}_{srchlt} = \frac{\overline{\text{within-class variability}} - \overline{\text{across-class variability}}}{\sigma\,(\text{across-class variability})}$$

Poster and appendix online

Picture – word contrast$_{srchlt}$



Picture – word contrast$_{srchlt}$ for single objects

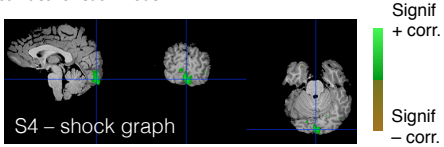


Computer vision (CV) models of voxel object encoding

- Learned map *M* from CV-based features *f* to voxel responses *v*,

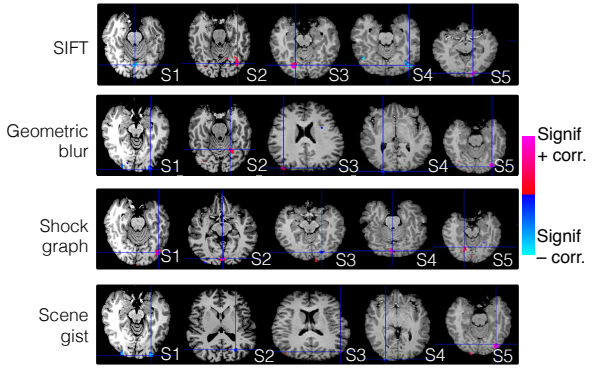
$$\hat{v} = Mf$$

- Compared match between predicted $\hat{v}$ and measured *v* voxel activities for each model



- Compared representational dissimilarity matrices (RDM) from voxel and CV image encodings (cf, Kriegeskorte 2008)

RDM(i,j) difference in voxel responses for stimuli *i* and *j*



Discussion

- In ventral pathway, coding more consistent for visual rather than semantic information
- Object contours (Shock graphs) predict BOLD activity in anterior visual regions
- Object sub-regions' features (SIFT, Geo. blur) predict some subjects' BOLD activity in distinct visual regions

References

M. Just, V. Cherkassky, S. Aryal, T. Mitchell "A Neurosemantic Theory of Concrete Noun Representation Based on the Underlying Brain Codes," *PLoS ONE*, 5(1), 2010.

N. Kriegeskorte, R. Goebel, P. Bandettini "Information-based functional brain mapping," *Neuroscience*, 103(10), 2006.

N, Kriegeskorte, M. Mur, P Bandettini "Representational similarity analysis – connecting the branches of systems neuroscience." *Frontiers in Sys Neurosci*, 2(4), 2008.

Acknowledgments

Funded by NSF IGERT, R.K. Mellon Foundation, NIH EUREKA Award #1R01MH084195-01, and the Temporal Dynamic of Learning Center at UCSD (NSF Science of Learning Center SBE-0542013)