Multi-Voxel Pattern Analysis of Noun and Verb Differences in Visual and Ventral Temporal Cortex

Christine Boylan [1,2] , John Trueswell [1] , Sharon L. Thompson-Schill [1,2]

(1) Department of Psychology, University of Pennsylvania

(2) Center for Cognitive Neuroscience , University of Pennsylvania

chboylan@sas.upenn.edu

Introduction

Recent evidence suggests that a probabilistic relationship exists between the phonological/orthographic form of a word and its lexical-syntactic category (specifically in nouns vs. verbs) [1]. Moreover, this so-called form typicality of a word with respect to its syntactic category has been found to modulate the M100 visual response in MEG [2]. Dikker et al. (2010) hypothesized that predictions about upcoming lexical-syntactic categories (e.g. nouns vs. verbs) give rise to form-based estimates in sensory cortex. We tested this hypothesis by conducting multi-voxel pattern analysis (MVPA) over V1 and left ventral temporal (VT) cortex (including the so-called "visual word form area") when subjects were predicting, *but crucially not viewing*, nouns and verbs. This allowed us to investigate prediction effects in these two ROIs without any bottom-up orthographic input. We used the brain data to classify the prediction of nouns vs. verbs in both sentence and non-sentence contexts.

Methods and Materials

Experiment I: Sentence Context Cue

Experiment II: Individual Word Cue

Subjects: Ten (Expt 1) and four (Expt 2) undergraduates at the University of Pennsylvania, all right-handed native speakers of English.

Experiment I

Materials: Sentences had low lexical cloze probability (mean cloze probability = 2.8%, range: 1.3%-29.3%) but high selectivity (100%) for either noun or verb completions (48 noun-type, 48 verb-type sentences). Sentence completions were normed over 75 subjects.

Noun1: (24 sentences)

Wh	V_{aux}	NP	PP______?
Where	*was*	*the woman*	*for the ______?*

Noun2: (24)

Wh	V_{aux}	NP	VP______?
When	*did*	*the janitor*	*mention the ______?*

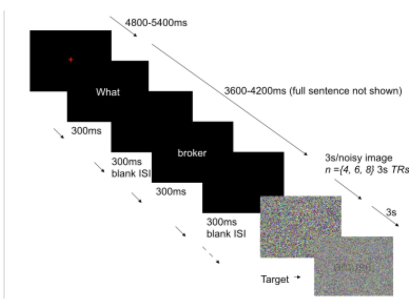
Verb1: (24)

Wh	NP	V_{aux}	NP	VP______?
Which	*budget*	*was*	*the mechanic*	*permitted to ______?*

Verb2 : (24)

Wh	NP	V_{aux}	NP	VP______?
What	*crib*	*did*	*the broker*	*plan to ______?*

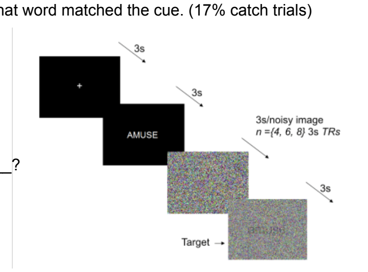
Task: Instead of seeing the sentence-final word immediately, subjects searched for a word in a series of noisy images. When a word was presented at subject's threshold, subject indicated with a button press whether that word was an appropriate completion to the sentence. (17% catch trials)



Experiment II

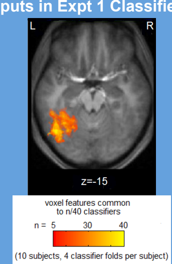
Materials: Tokens: two typical nouns (*bible, movie*); two typical verbs (*adopt, amuse)* (taken from [1,2]).

Task: Subject was cued to look for a word in noise. When a word was presented at subject's threshold, subject indicated with a button press whether that word matched the cue. (17% catch trials)

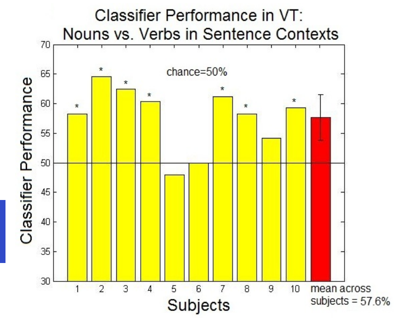


Analyzing only those volumes collected when *subjects predicted a word but saw pure noise*, we implemented a simple neural network with an input layer consisting of those 200 voxels best accounting for the variance between noun and verb conditions in each ROI in each run (taking best F scores) [3]. This was trained on three runs using a conjugate gradient descent backpropagation algorithm [4], and tested on a fourth run in a leave-one-out 4-fold cross-validation procedure. Within-subject classifiers were sought for sentence-context noun-vs.-verb prediction (experiment 1) and individual form-typical word prediction (experiment 2).

Density Map of VT Voxel Inputs in Expt 1 Classifiers

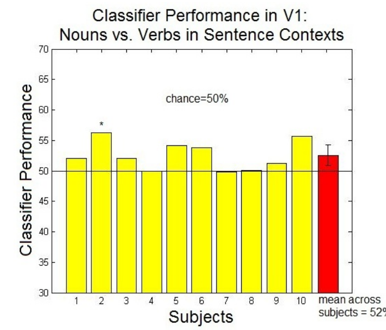


Experiment 1: Classification of Sentence-Context Prediction



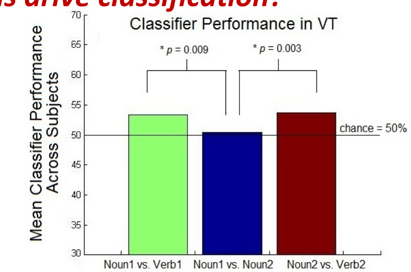
Mean classification performance of nouns vs. verbs in sentence contexts was significant for 7 (of 10) subjects in VT (mean across subjects: 58%, chance = 50%) and for 1 (of 10) in V1 (mean: 52%).

* Significance testing within subjects: Regressors were scrambled such that the new set had the same number of trials and conditions per run as the original regressor set. The regressor set was randomly permuted 1,000x to simulate the null distribution, and 1,000 cross-validation iterations were performed. (α =.05)



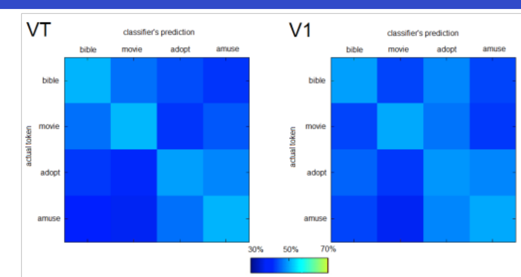
Does number of content words drive classification?

To test whether the number of content words in each sentence condition might account for classification accuracy above, we tested classification of Noun1- vs. Noun2-type sentences. Classifier performance was not significantly above chance. To check whether low power might mask an effect here (24 sentences rather than 48 in each condition), we also tested Noun1 vs. Verb1 and Noun2 vs. Verb2, classification pairs matched in power to Noun1 vs. Noun2. These cross-category classifications were reliably above chance despite decreased power.



Experiment 2: Classification of Individual Word Prediction

Individual word classifier consistently predicted the correct token more often than an incorrect token in both V1 and VT. Moreover, in VT a given noun was more often confused with another noun than with a verb, and vice-versa (chi-squared test , $p< 0.05$ for VT, not V1). See pairwise confusion matrix to right (chance =50%).
Mean classification in VT in 4-way classification: 29.3%; mean classification in V1: 27.8% (chance = 25%).



Conclusion

The sentence-context prediction results (experiment 1) suggest that syntactic cues are sufficient to drive top-down predictions of word form features in VT. The within-category confusability in VT of the individual word predictions (experiment 2), for which lexical category was not necessary to predict the cued word form, suggests that retrieval of lexical category information may be automatic during word prediction.

References
1. Farmer, T. A., Christiansen, M. H., and Monaghan, P., 2006. P Natl Acad Sci USA 10332, 12203-12208.
2. Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L., 2010. Psychol Sci 1;215: 629-34.
3. McDuff, S.G.R., Frankel, H.C., Norman, K.A. 2009. J Neurosci 292: 508-516.
4. Polyn, S.M., Natu, V.S., Cohen, J.D., Norman, K.A., 2005. Science 310: 1963–1966.

Acknowledgments

Thanks to the labs of Sharon Thompson-Schill and John Trueswell for their advice and input. Special thanks to Marc Coutanche for aiding in design and analyses.
This research was supported by grants R01MH70850, R01EY021717 (Sharon Thompson-Schill) and NNC P30 NS45839 (Center for Functional Neuroimaging).