Right-Lateralization of Verbal Collocations

Shohini Bhattasali Murielle Fabre John Hale
Cornell University

24th Architectures and Mechanisms of Language Processing conference, September 2018

Introduction

Natural language comprehension relies on at least two cognitive processes:

- Retrieval of memorized elements
- Structural composition

Frequent collocations like *break the ice*, *boa constrictor*, *safe and sound*, *see to it*, *in spite of* can help us address the neural bases of these processes. These collocations, often known as Multiword Expressions, form a heterogeneous family of word clusters.

Questions

1. Are the differences between the grammatical categories of MWEs observable at the cerebral level?
2. Does processing of verbal MWEs implicate separate brain areas from non-verbal MWEs?

Data Collection

Participants (n=51) were college-aged, right-handed, native English speakers.
Listened to a spoken recitation of *The Little Prince* for 1 hour and 38 minutes across nine separate sections; 15,388 words in total.
Comprehension was confirmed through multiple-choice questions at the end of each section (90% accuracy, SD = 3.7%).

Multiword Expressions

742 MWEs were identified in the dataset through a transition-based MWE analyzer (Al Saied et al., 2017) trained on Children's Book Test dataset (Hill et al., 2015).
Stanford POS tagger & NLTK POS tagger were used to annotate the grammatical categories of MWEs' words (Bird and Loper, 2004; Manning et al., 2014).
Presence/absence of verbal expression yielded two categories of MWEs (i.e. 56% verbal vs. 44% non-verbal).

Group-level Results

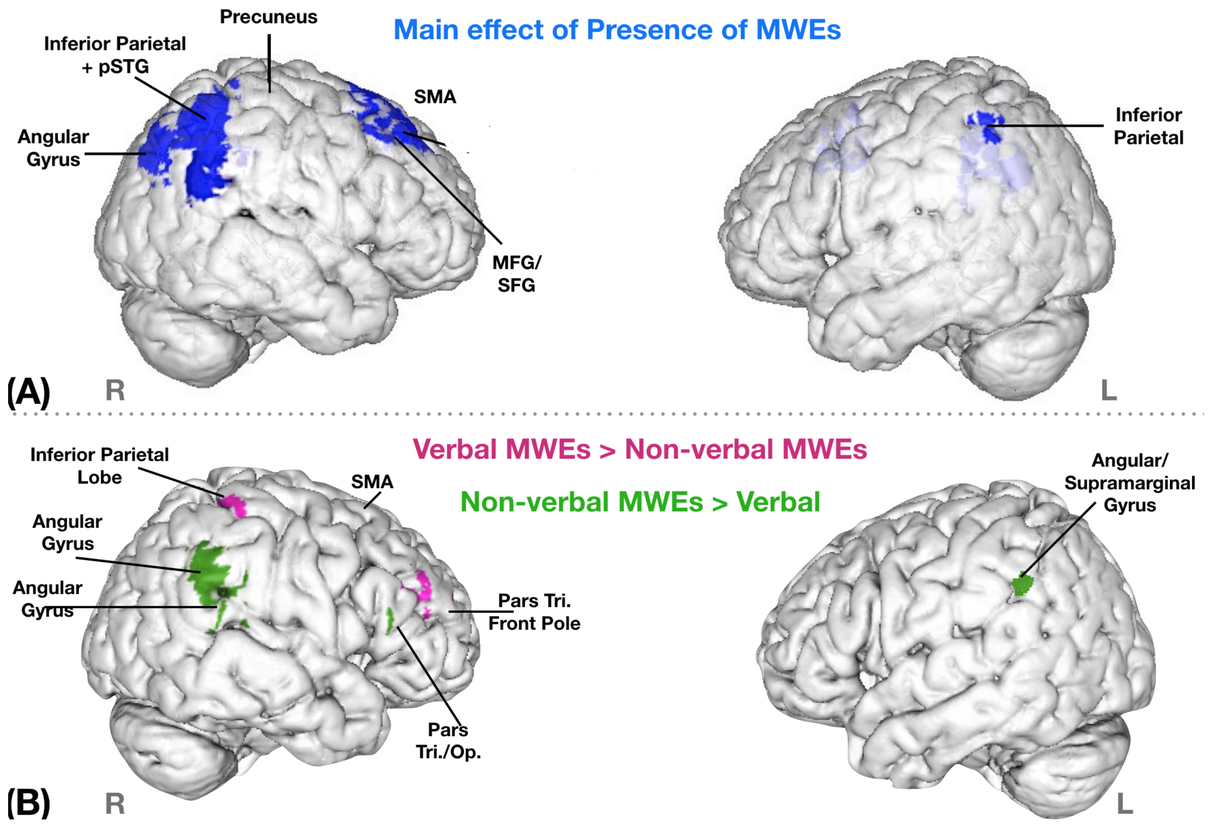

Figure 1: (A): Whole-brain effect for the presence of MWEs in blue. (B): Contrast images with significant clusters for [Verbal MWEs > Non-verbal MWEs] in pink and [Non-verbal MWEs > Verbal MWEs] in green. (FWE $p < 0.05$)

fMRI Analysis

Preprocessing was carried out with AFNI version 16 and ME-ICA v3.2 (Kundu et al., 2011).

- ME-ICA denoises T2* signal using ICA into BOLD and noise components from physiology, motion, scanner artifacts.

MWE predictors convolved with SPM12's canonical HRF, regressed against observed BOLD signal during passive story listening.

To account for sentence-level compositional processes, a regressor formalizing syntactic structure building included (Hale, 2014).

GLM analysis includes four regressors of non-interest: word offset, frequency, pitch, intensity.

Conclusion

1. Main effect of MWEs implicate areas associated with lexical semantic network (Binder et al., 2009).
2. Verbal and non-verbal MWEs show spatially distinct patterns of activation.

Discussion

- Bilateral Angular Gyrus/Supramarginal Gyrus is sensitive to co-occurrence frequency of word combinations (Graves et al., 2010) & argument structure (Meltzer-Asscher et al., 2013).
- While not part of the language network, Precuneus implicated in memory and naming tasks (Halsband et al., 2002; Platel et al., 1997).
- Results suggest that verb-argument selectional relations in verbal collocations involve right hemisphere activity in Broca's area and IPL (Thompson et al., 2007).
- Surprisingly right-lateralized network: semantic richness of MWEs (Price et al., 2015) & naturalistic stimuli (Wehbe et al., 2014).

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1607441.

Selected References

Al Saied, H., Constant, M., & Candito, M. (2017). The ATILF-LLF system for PARSEME shared task: A transition-based verbal multiword expression tagger. In *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)* (pp. 127–132).
Hale, J. T. (2014). *Automaton theories of human sentence comprehension.* CSLI Publications.
Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M., & Bandettini, P. A. (2012). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage, 60*(3), 1759–1770.