Modeling Conventionalization and Predictability in Multiword Expressions at the Brain Level
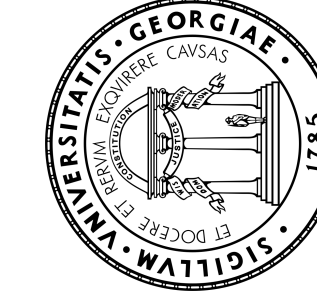
Shohini Bhattasali [1] Murielle Popa-Fabre [2] John Hale [3] Christophe Pallier [4]

[1]University of Maryland [2]INRIA – University of Paris [3]University of Georgia [4]CEA Inserm Cognitive Neuroimaging Unit

SCiL 2020

Question

Which gradient metric is psychologically plausible for processing these noncompositional expressions during language comprehension?

MWE Examples: break the ice, boa constrictor, safe and sound, see to it, in spite of

Key Ideas

- Apply computational linguistics metrics in a naturalistic listening fMRI study
- r^2 Model Comparison, instead of GLM localization
- Dice's Coefficent is a better predictor of neural activation for processing MWEs

Data Collection

51 native English speakers listened to *The Little Prince* for 1 hr & 38 mins across nine separate sections; 15,388 words in total.

Comprehension confirmed through multiple-choice questions (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).

Association Measures calculated using corpus frequency counts from COCA:

1 Pointwise Mutual Information (PMI):

$$\text{PMI} = log_2 \frac{c(w_n^1)}{E(w_n^1)} \qquad (1)$$

2 Dice's Coefficient:

$$\text{Dice} = \frac{n \times c(w_n^1)}{\Sigma_{i=1}^{n} c(w_i)} \qquad (2)$$

Group-level Results

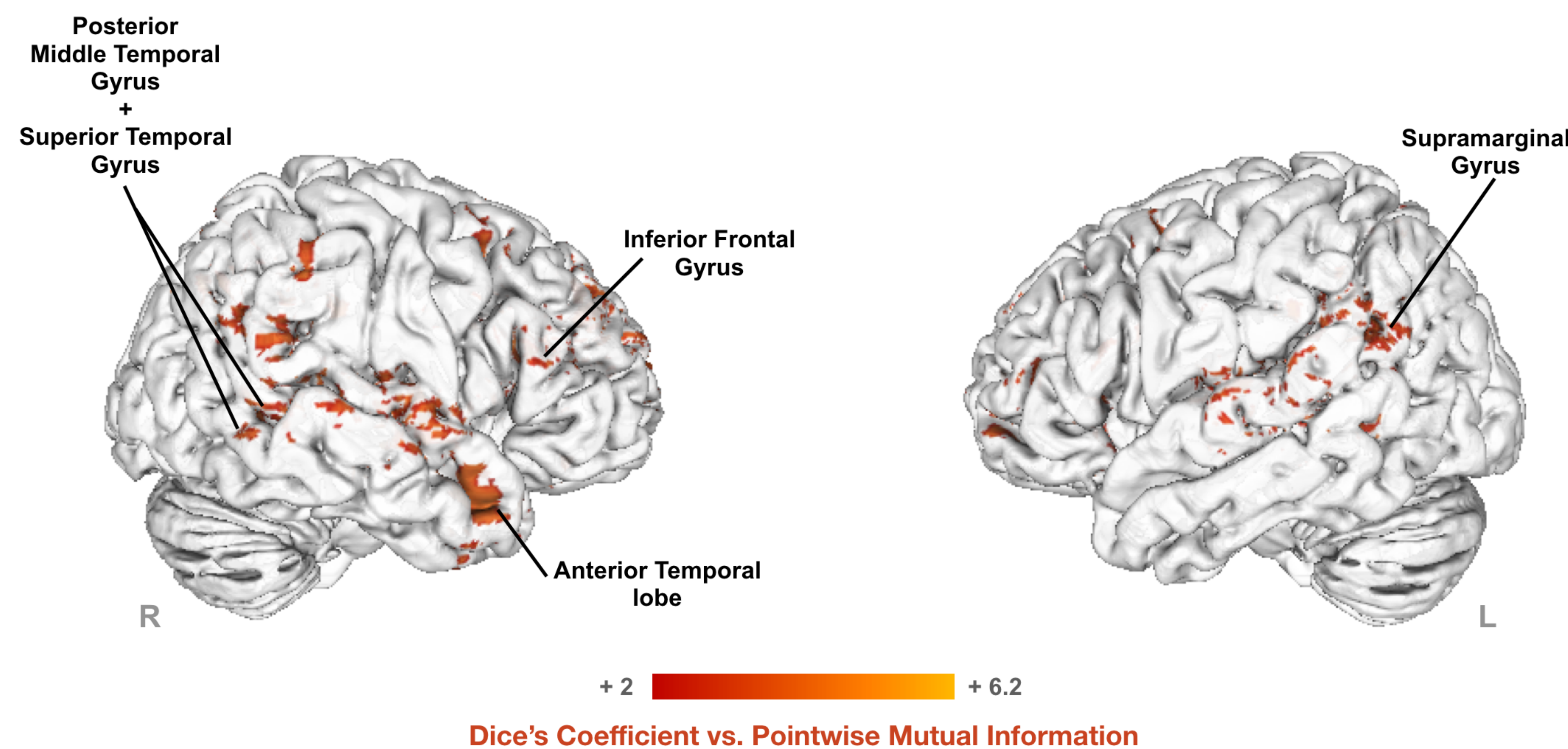**Dice's Coefficient vs. Pointwise Mutual Information**

+ 2   + 6.2

Figure 1:Z-map showing regions having a significant effect for Dice's coefficient versus Pointwise Mutual Information after Bonferroni correction with p < 0.05
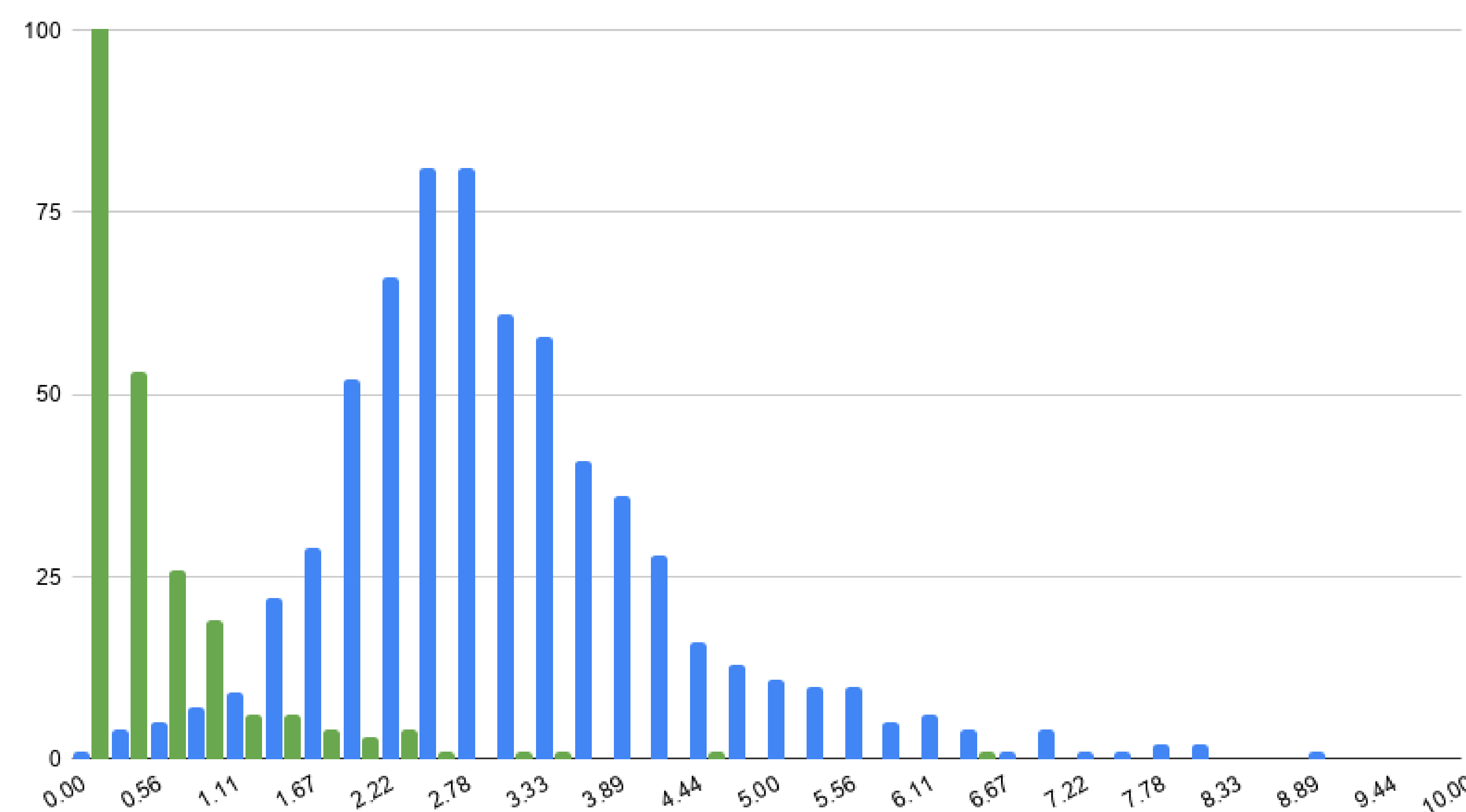

Figure 2:Graph comparing Pointwise Mutual Information (in blue) with Dice's Coefficient (in green).

fMRI Analysis

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

GLM Analysis:

- MWE predictors convolved with SPM12's canonical HRF, regressed against observed BOLD signal.
- Includes four regressors of non-interest: word offset, frequency, pitch, intensity

r^2 Model Comparison: For every subject, we compute how much the inclusion of each variable of interest (i.e. Dice and PMI) increases the cross-validated r^2.

Conclusion

- Dice's Coefficient, formalizing the degree of predictability, is a better predictor of cerebral activation for processing MWEs.
- Indicates that Dice is a more cognitively plausible computational metric.
- Previous work has shown gradient nature of MWEs (Bhattasali et al., 2018); results here illustrate how gradient they are.

Acknowledgments