Background

- Within cognitive categories, the accessibility of items is not equal. This can be measured using the frequency of production of items.
- Constructions are emergent verb-centered categories. Some verbs within each construction are produced more frequently than others. What predicts their frequency of production?

Method

**Computational Model**
- Has been used for simulating human categorization (Anderson, 1991), L1 (Alishahi & Stevenson, 2008) and L2 construction learning (Matusevych et al., submitted).
- A random example instance and its processing.

**Example Instance:**
- \( \text{Learning from input: find the best construction for an instance } I. \)
- \( C_{\text{args}}(I) = \text{argmax } P(C|I) \)
- Prior: entrenchment; conditional: similarity.

\[ P(C) = \frac{|C|}{N+1} \]

\[ P(I|C) = \prod_{v \in I} P(F(v|C)) \]

**Simulation Experiment**
- 100 simulations with individual inputs sampled from the data set.
- Given a test instance, produce a list of verbs.

**Variables Definitions**
- Outcome variables:
  1. Frequency \( PF(v,c) \)
  2. Probability \( PP(v,c) \)
- Predictor variables:
  1. \( F(v,c) \) and \( F(v) \) directly available
  2. \( \Delta P(v,c) = P(v|c) - P(v|\neg c) \)
  3. \( A(v,c) = P(v|c) = \frac{F(v,c)}{F(c)} \)
  4. \( \text{Prt}(v,c) = \sum_{v \in I} \frac{|M \cap M_j|}{|c|} \)

Results

**Replicating Original Experiments**
- \( PF(v) \sim F(v,c) \times \Delta P(v,c) \times \text{Prt}(v,c) \)
- Relative importance of predictors: \( F(v,c), \Delta P(v,c), \text{and Prt}(v,c) \).

**Fixing Methodological Problems**
- \( PP(v) \sim F(v,c) \times \Delta P(v,c) \times \text{Prt}(v,c) \)

**Refining Prediction Model**
- \( I \): marginal frequency
- \( II \): association
- \( III \): semantics

**Conclusions**
- Computational modeling on a different data shows similar results.
- The predictive power of variables may be lower than originally thought.
- Marginal frequency is an independent factor.
- Parallel use of two association measures leads to collinearity.
- Attraction is the best predictor in our data.
- The impact of prototypical is lower than expected.

Study Objectives

I. Address the two issues.
II. Refine the prediction model.

Computational modeling!