Across languages, adjectives show preferences for some orderings over others. “the large wooden table” > “the wooden large table” “the beautiful old house” > “the old beautiful table”

Generalizations include adjectives
- Inherency (Whorf 1945) and Specificity to the noun (Sweet 1898, Ziff 1960)
- Absoluteness (Sproat & Shih 1991) and Subjectivity (Hetzron 1978)

but there is no consensus on explanation.

**Hypothesis:**

Across languages, mutual information with noun impacts adjective ordering, as expected under assumptions of memory limitations and pressure to minimize prediction error.

**Conclusion:**

Across languages, mutual information with noun impacts adjective ordering, as expected under assumptions of memory limitations and pressure to minimize prediction error.

**Mutual Information**

MI(Adj,Noun) = log P(Noun|Adj) - log P(Noun)

Measures collocation, plausible measure of specificity.

Prediction: Adjectives with higher MI with noun are closer to it

**Experiment I**

Hypothesis: Higher MI ~ Closer to Noun

MI with Noun (in bits)

<table>
<thead>
<tr>
<th>Language</th>
<th>English</th>
<th>Chinese</th>
<th>Russian</th>
<th>Arabic</th>
<th>Spanish</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI</td>
<td>4.2</td>
<td>4.1</td>
<td>3.8</td>
<td>4.3</td>
<td>2.9</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Across languages, adjectives with higher MI with noun occur closer to noun.

**Experiment II**

**Question:** Do differences in $A_1$, $A_2$ Mutual Information with the noun affect ease of incremental prediction for powerful statistical models of language?

**Prediction:** Lower aggregate surprisal on real ordering than reverse orders

Use recurrent neural network models. Order preferences were eliminated from training data.

We compare surprisal on real and reverse orderings:

- Average per-word surprisals (schematic)
- 1.0 bits $p < 0.0001$

For statistical model of language, real orderings maximize ease of incremental prediction.

**Experiment III**

Motivation:
Scontras et al. (2017) show: The more subjective an adjective, the farther away from the noun it occurs.

**Question:** Do MI and subjectivity contribute to ordering preferences independently?

**Model Comparison with MI-only model:**

MI+Subjective is better than MI alone

**Conclusion:**
Across languages, mutual information with noun impacts adjective ordering, as expected under assumptions of memory limitations and pressure to minimize prediction error.