A large-scale deconvolutional study of predictability and frequency effects in naturalistic reading

Cory Shain
Ohio State

shain.3@osu.edu

---

### Experimental Design

- **Evaluation on 3 large corpora containing over 1M events total:**
  - Natural Stories, self-paced reading [7]
  - Dundee, eye-tracking [13]
  - UCL, eye-tracking [6]
- **Deconvolutional time series regression** [19]
- **Costs:** Sentence position, document position, word rate, word length, saccade length, whether the previous word was fixed
- **Predictors of interest:** unigram log probability, 5-gram surprisal
- **Response:** Log-ms (go-past for eye-tracking)
- **By-subject random intercepts, slopes, and impulse response parameters

### Results

<table>
<thead>
<tr>
<th>Corpus</th>
<th>SentPos</th>
<th>Trial</th>
<th>Rate</th>
<th>WordLen</th>
<th>SacLen</th>
<th>PrevFix</th>
<th>Unigram</th>
<th>5-gram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural Stories</td>
<td>0.0098</td>
<td>-0.0216</td>
<td>-0.3069</td>
<td>—</td>
<td>—</td>
<td>0.0158</td>
<td>-0.0018</td>
<td>0.0174</td>
</tr>
<tr>
<td>Dundee</td>
<td>-0.0085</td>
<td>-0.0052</td>
<td>-0.0277</td>
<td>0.0068</td>
<td>-0.0021</td>
<td>-0.0178</td>
<td>-0.0067</td>
<td>0.0117</td>
</tr>
<tr>
<td>UCL</td>
<td>0.0524</td>
<td>-0.1330</td>
<td>0.0023</td>
<td>0.0221</td>
<td>0.0778</td>
<td>0.0005</td>
<td>0.0184</td>
<td></td>
</tr>
</tbody>
</table>

- **Comparison:** 5-gram only vs. baseline
- **Unigram only vs. baseline** 0.0001***
- **5-gram + Unigram vs. Unigram-only** 0.0001***
- **5-gram + Unigram vs. 5-gram-only** 0.1440

- **Post-hoc permutation testing results on out-of-sample data.**

- **Main result:** Significant effect of **frequency** over **predictability**, but not vice versa, consistent with **No**.

- **Frequency effects are larger magnitude than frequency effects.**

### Conclusion

**Results support No:** no evidence of separable effects of frequency and predictability.

- **Finding is at odds with constructed experiments.**
- **Possible explanations:**
  - Frequency effects may exist in naturalistic reading but are too small to be detected.
  - Constructed stimuli may introduce confounds:
    - Atypical word distributions
    - Lack of context
    - Suspension of normal communicative function
    - Cloze models may be too coarse, allowing frequency predictors to capture residual variance due to predictability

---

### References