**Background**

Comprehenders incur a cost while reading an anaphor that mismatches the gender typicality of its antecedent\(^1\)-\(^3\). Why?

- One answer: Surprisal. Comprehenders incrementally generate expectations of a word's probability in context\(^4\).
- Mismatch cost: unexpected word form given referential context.
- Prediction: The mismatch effect should be fully continuous across the range of gender typicalities.

**Current Work**

We use eye-tracking to test for graded effects of gender typicality, to determine i) whether mismatch effects are graded/continuous and ii) the time-course of these effects.

**Materials: Gender norming**

Participants (N = 42, Mage = 19) estimated the extent to which a role noun associated with women and men (i.e., the ratio) on an 11-point rating scale.

<table>
<thead>
<tr>
<th>Male typical</th>
<th>Female typical</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>0.9</td>
<td>1.0</td>
</tr>
</tbody>
</table>

**Experiment: Eye-tracking**

**Method:**
- Eye tracking with comprehension question task (N = 48; M\(_\text{age} = 19\)).
- 45 critical items constructed in a 1x3 design as in (1).
- 2 items created for each level of gender typicality from .15 male to .85 male (N\(_\text{run} = 30\)).
- 15 items created with definitional gender.
- Combined with 55 fillers.

1. The X hurt **herself/they** critically/spillover

**Design:**
- 1x3 within-subjects manipulation
- Antecedent nouns were placed in transitive sentence frame, with one of three direct object reflexives: **himself/herself/they**.
- Reflexives selected because structural constraints minimize role of antecedent identification in processing the pronoun. Thus, a clearer test case to evaluate the impact of graded gender typicality.

**Analysis:**
- Log transformed RTs were analyzed using LME model with **typicality** as a fixed-effects factor and subjects and items as random factors.
- Each pronoun analyzed separately.
- I adopt the convention that a t-value greater than 2 presents a significant effect\(^6\).

**Results**

- First fixation: Significant effects of typicality at the critical region for **himself** and **herself** but no effect on **themselves**.
- First pass: Significant effects of typicality for **himself** and **herself** but no effect on **themselves** at both the critical and spillover regions.
- Go-past: Significant effects of typicality for **himself** and **herself** at both the critical and spillover regions. Crucially, the LME model revealed significant effect for typicality for **themselves** when used a quadratic predictor.

**References**


**Discussion**

- First fixation and first pass show continuous effect of gender typicality on both **himself** and **herself**.
- Sharp increase in reading times for mismatched definitional nouns for **himself/herself**, especially in go-past times.
- Late effect of typicality on **themselves**: reading time scales with typicality.

**Take away**

- Reflexives provide evidence that comprehenders generate graded expectations about gender typicality (supporting the surprisal view).
- Effects arise as early as first fixation\(^7\).
- However, recognition of gender-neutral orientation of **themselves** comes only at a delay.

**Acknowledgements**

I would like to thank Kyle Johnson and the members of the Cognitive Science Lab at UMass Amherst and Child Language Lab at NYU for their feedback.

**Take away**

- Reflexives provide evidence that comprehenders generate graded expectations about gender typicality (supporting the surprisal view).
- Effects arise as early as first fixation\(^7\).
- However, recognition of gender-neutral orientation of **themselves** comes only at a delay.