Grammatical and Coherence-Driven Biases in Pronoun Interpretation

Hannah Rohde
Northwestern University

Andrew Kehler
UC San Diego
Interpreting ambiguous pronouns

- Morphosyntactic factors: grammatical-role biases influence interpretation (e.g., Kameyama 1996)

  (1) **John** kicked Bill. Mary told **him** to go home.
  (2) **Bill** was kicked by John. Mary told **him** to go home.

- Coherence-driven factors: interpretation is a side effect of general inferencing used in establishing a coherent discourse (Hobbs 1979, Kehler 2002)

  (3) The **city council** denied the **demonstrators** a permit...
      a. ... because **they** feared violence
      b. ... because **they** advocated violence

  (Winograd 1972)
This talk

**Puzzle:** Evidence that pronoun interpretation is sensitive to grammatical role biases and coherence-driven factors

**The question:** What type of model can capture these facts?

Outline

- Evidence for coherence-driven biases (Rohde et al. 2006, 2007)
- Evidence for grammatical-role biases (Stevenson et al. 1994)
- Story continuation experiment
- Bayesian model of pronoun interpretation
- Conclusions
Evidence for coherence-driven biases (Rohde et al. 2006, 2007)

Pronoun interpretation biases vary across coherence relations

(4) John handed a book to Bill. He ________________.

(5) John passed the comic to Bill. He took it and opened it.

(6) John passed the comic to Bill. He did so carefully.

(7) John passed the comic to Bill. He wanted Bill to read it.

(8) John passed the comic to Bill. He accidentally hit Bill.

(9) John passed the comic to Bill. He thanked John.
Coherence breakdown (Rohde et al. 2006, 2007)

John passed the comic to Bill. He ________________.

\[ p(\text{ref}) = \sum_{\text{coh}} p(\text{ref} | \text{coh}) \times p(\text{coh}) \]

→ Results show role of coherence establishment in interpretation
Interpretation/production asymmetry

- **Story continuations** (Stevenson et al. 1994)

  (10) John passed the comic to Bill. He _____________.
  (11) John passed the comic to Bill. _______________

- **Results** (see also Arnold 2001)

  - INTERPRETATION: Pronoun prompt in (10) yields 50/50 Source/Goal interpretation -- participants were equally likely to interpret pronoun to refer to Goal as to Source
  - PRODUCTION: No-pronoun prompt in (11) reveals subject production bias for pronouns... participants were more likely to produce a pronoun when re-mentioning subject/Source and a name when re-mentioning non-subject/Goal

  - Why is this a problem? Hobbsian coherence-driven model
    - pronouns are unbound variables
    - no misalignment between interpretation/production is expected
Prediction for interacting biases

- **Prediction:** bias to produce a pronoun when referring back to subject referent may influence coherence establishment

  \[ p(\text{coh}) = \sum_{\text{ref}} p(\text{coh}|\text{ref}) \times p(\text{ref}) \]

→ **Production influences interpretation:** based on production bias in Stevenson et al., \( p(\text{ref}=\text{subj}) \) is higher when pronoun is present

→ **Which in turn influences coherence:** probability that upcoming coherence relation is subject-biased (Source-biased) is higher when pronoun is present, even one whose referent is fully ambiguous

→ Are there more Source-biased relations with pronoun prompt?

- **Method:** elicit story continuations, categorizing referents of referring expressions and categorizing passages for coherence

  Pronoun prompt: John passed the comic to Bill. He _____.
  No-Pronoun prompt: John passed the comic to Bill. _____.
Story continuation experiment

- Participants: 52 monolingual English speakers
- Task: instructed to write a natural continuation
- Stimuli: vary prompt type (pronoun/no-pronoun)
- Evaluation: two judges assess coherence and coreference
Results: choice of referring expression

- Continuations elicited with no-pronoun prompt confirm subject bias in participants’ production of pronouns

No-Pronoun Prompt: John passed the comic to Bill. ____.
Results: coreference

- Presence of pronoun yields more references to the subject of the previous sentence (Source references)

Pronoun prompt: John passed the comic to Bill. He _____.
No-Pronoun prompt: John passed the comic to Bill. _____.

Error Bars: +/- 1 SE
Results: coherence

- Compare proportions of two most frequent and opposite-biased coherence relations (Source-biased Elaborations & Goal-biased Occasions) and test for effect of prompt type

- We find more Elaborations with pronoun prompt and more Occasions with no-pronoun prompt

→ Presence of ambiguous pronoun shifts coherence
→ Confirms prediction regarding interaction of coherence-driven biases and grammatical-role biases
Accounting for the asymmetry

- Bayes rule for pronoun interpretation

\[
p(\text{ref} \mid \text{pro}) = \frac{p(\text{ref}) \cdot p(\text{pro} \mid \text{ref})}{p(\text{pro})}
\]

Prior for referent in a particular context

Pronoun interpretation: Stevenson et al.'s result that probability is 50/50 of interpreting a pronoun as subject-referring

Overall probability of using a pronoun

Grammatical-role production bias: Stevenson et al.'s result that probability of producing a pronoun is high when referent is the subject
Accounting for the asymmetry

- Case study to see asymmetry: subject #14

\[ p(\text{ref} \mid \text{pro}) = \frac{p(\text{ref}) \times p(\text{pro} \mid \text{ref})}{p(\text{pro})} \]

<table>
<thead>
<tr>
<th>Interpretation bias (pro prompt)</th>
<th>Production biases (no-pro prompt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p(\text{ref} = \text{Source} \mid \text{pro}) = 0.63 )</td>
<td>( p(\text{ref} = \text{Source}) = 0.33 )</td>
</tr>
<tr>
<td></td>
<td>( p(\text{pro} \mid \text{ref} = \text{Source}) = 1.0 )</td>
</tr>
<tr>
<td></td>
<td>( p(\text{pro} \mid \text{ref} = \text{Goal}) = 0.25 )</td>
</tr>
<tr>
<td>( p(\text{ref} \mid \text{pro}) = 0.63 )</td>
<td>( \frac{p(\text{ref}) \times p(\text{pro} \mid \text{ref})}{p(\text{pro})} = 0.67 )</td>
</tr>
</tbody>
</table>

- Rate of Source coreference: correlated across prompt types

\[ R^2 = 0.266, F(1,100) = 37.65, p < 0.001; R^2 = 0.404, F(1,71) = 49.73, p < 0.001 \]
Biases regarding upcoming coherence relations are dependent on the probability that a particular referent has been mentioned.

\[
p(\text{coh}) = \sum_{\text{ref}} p(\text{coh} | \text{ref}) \times p(\text{ref})
\]

Probability of next mention is dependent on the presence of a pronoun:

\[
p(\text{ref} | \text{pro}) = \frac{p(\text{ref}) \times p(\text{pro} | \text{ref})}{p(\text{pro})}
\]
Summary

- Introduced a model to capture both coherence-driven and grammatical-role biases
  - crucially without simply enumerating factors but rather showing when and how particular factors come into play
- We explain the apparent asymmetry between pronoun interpretation and production
- Our model correctly predicts that the occurrence of a pronoun -- even a fully ambiguous one -- will change expectations about how the discourse will continue
Acknowledgements

Thanks to research assistant Erica Gold.

Thanks also to Roger Levy, Brady Clark, Masaya Yoshida, and the Northwestern Synax & Meaning Lab for their comments on earlier drafts.

This work was supported by NIH grant T32-DC000041 to the Center for Research in Language at the University of California, San Diego.

Thank you!