



The Bidirectional Influence between Coherence Establishment and Pronoun Interpretation

Hannah Rohde & Andrew Kehler

Abstract

Problem: Previous work finds that pronoun interpretation in transfer-of-possession contexts is sensitive to coherence-driven biases, but that work does not address a mystery that has also been reported for such contexts: an asymmetry between pronoun production and interpretation.

Proposal: To explain the asymmetry, we introduce a Bayesian model of pronoun interpretation that distinguishes the primary factors (e.g., coherence relations) that generate expectations about who will be mentioned next (p(referent)) from those factors (e.g., subjecthood) that primarily determine whether a speaker will choose to use a pronoun instead of a name (p(pronoun | referent)). We predict more subject-biased coherence relations than non-subject-biased relations in a pronoun-prompt story-continuation condition.

Results: We find that the rate at which a pronoun is interpreted to refer to the subject reflects the likelihood that the subject referent will be mentioned again (regardless of referring expression) and the likelihood that a speaker will use a pronoun. Crucially, we find that, in contexts with an ambiguous pronoun prompt, the mere presence of a pronoun, even if its referent is ambiguous, can change the distribution of coherence relations.

1. Goals

- Distinguish factors that influence comprehenders' expectations for next mention from those that influence speakers' choice of referring expression
- Test whether ambiguous pronouns influence expectations about discourse direction given a production bias linking pronouns and subjects

2. Previous work: pronoun interpretation

Transfer of possession contexts

(1) John_{SOURCE} passed the comic to Bill_{GOAL}. He _____.

Stevenson, Crawley, & Kleinman (1994): interpretation biases compete

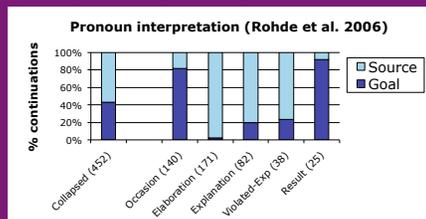
- Thematic role bias (favoring Goal over Source)
- Subject assignment bias (favoring subject Source)

RESULT: participants interpret pronoun 50/50 to Source/Goal

Rohde, Kehler, & Elman (2006): interpretation reflects discourse biases

- Biases about where discourse is going (coherence relations)
- Biases about who is likely to be mentioned next (given discourse direction)

RESULT: interpretation depends on operative coherence relation



→ This result does not cast any light on the interpretation/production asymmetry.

3. Interpretation/Production asymmetry

Prompt/no-prompt conditions reveal apparent asymmetry (Stevenson et al. 1994; Arnold 2001)

(2) a. John passed the comic to Bill. He _____
 b. John passed the comic to Bill. _____

Choice of next mention

- (2a) – pronoun interpretation 50/50 to Source/Goal
- (2b) – strong bias to refer to Goal as next mention

Choice of referring expression

- (2b) – pronoun production biased to subject Source
- (2b) – name production biased to non-subject Goal

Stevenson et al.: asymmetry provides evidence for subject assignment bias that competes with thematic role bias

→ Claim for 2 biases: an expectation for a subsequent mention of a referent, p(referent), and an expectation about the form of referring expression that the speaker would use to mention that referent, p(proreferent).

4. Proposal: Bayesian model

$$p(\text{Goal}) \quad p(\text{Goal} | \text{pronoun}) = \frac{p(\text{pronoun} | \text{Goal}) * p(\text{Goal})}{p(\text{pronoun})}$$

(2b) as per Arnold

$$= .856 \quad \frac{30 * 149}{149 * 174} = .612$$

Claims:

- p(Goal) reflects coherence biases (e.g. Occasions biased to Goal)
- p(pronoun | Goal) reflects production preferences (i.e. speakers are more likely to pronominalize a reference to subject)
- Prediction: presence of a pronoun in (2a) will yield more subject Source interpretations and more Source-biased coherence relations

$$p(\text{CR}) = p(\text{CR} | \text{Goal}) * p(\text{Goal}) + p(\text{CR} | \text{Source}) * p(\text{Source})$$

5. Story continuation experiment

Task: participants instructed to imagine a natural continuation to a context sentence and prompt

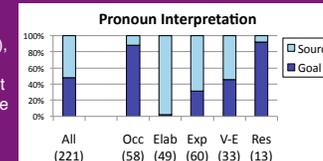
Materials: pronoun prompt (2a) and bare prompt (2b)

Judges: annotation for next mention, intended pronoun interpretation, and coherence relation

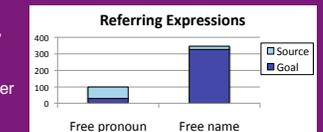
6. Coreference results

Replication:

- **Pronoun interpretation:** for (2a), pronoun is interpreted 50/50 to Source/Goal (as per Stevenson et al.), but biases differ by coherence relation (as per Rohde et al.)



- **Referring expressions:** for (2b), pronouns are used to refer to the Source, while names refer overwhelmingly to the Goal (as per Stevenson et al., and Arnold)

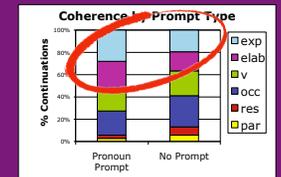


7. Coherence results

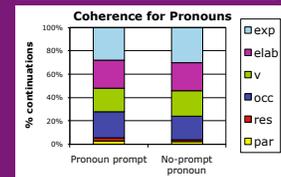
Prompt type: as predicted, the pronoun prompt (2a) yielded more Source-biased relations than did the bare prompt (2b)

Exp: $F_1(1,51)=13.373, p<0.001$
 $F_2(1,20)=15.218, p<0.001$

Elab: $F_1(1,51)=2.772, p<0.102$
 $F_2(1,20)=5.395, p<0.031$



Source-biased pronouns: as predicted, there was no significant difference in the distribution of coherence relations between (2a) and (2b) when only continuations with pronouns are considered



8. Summary: bidirectional influences

- **Coherence influences pronoun interpretation:** choice of next mention, p(Goal), reflects coherence-driven biases

- **Pronoun interpretation influences coherence establishment:** knowing that pronouns are more likely to be produced for a subject referent shifts coherence distribution, even when pronoun is ambiguous

References

Arnold, J. E. (2001) The effects of thematic roles on pronoun use and frequency of reference. *Discourse Processes* 31:2-137-162.
 Rohde, H., A. Kehler, J. L. Elman (2006) Event Structure and Discourse Biases in Pronoun Interpretation. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Vancouver, BC, Canada, July 26-29, 2006.
 Stevenson, R., R. Crawley, & D. Kleinman. (1994) Thematic roles, focusing and the representation of events. *Language and Cognitive Processes* 9.