Demanding an Explanation: Implicit Causality Biases in Discourse Interpretation
Hannah Rohde & Andrew Kehler

Abstract
Problem: Previous passage-completion studies report strong biases regarding who will be mentioned next following implicit causality (IC) verbs with a ‘because’ prompt. However, these biases are reduced/eliminated with a full-stop prompt.

1. Goal
To clarify the effects of IC biases on discourse interpretation by distinguishing (i) next mention biases and (ii) biases toward upcoming coherence relations.

2. Previous work on Implicit Causality
Passage completions: strong IC bias to particular referent with ‘because’ prompt (Caramazza, Grober, Garvey, Yates 1974; McKoon, Greene, Ratcliff 1993) (see also)

(1) a. IC-1: John annoyed Mary because ______________ (bias to NP1=John)
b. IC-2: John scolded Mary because ______________ (bias to NP2=Mary)
c. Non-IC: John babysat Mary because ______________ (mixed biases)

However, next-mention bias reduced/eliminated with full stop prompt (Au 1986; inter alia)

(2) a. IC-1: John annoyed Mary ______________ (bias to NP1=John)
b. IC-2: John scolded Mary ______________ (bias to NP2=Mary)
c. Non-IC: John babysat Mary ______________ (mixed biases)

What is role of ‘because’?
• Modifying salience of event participants directly (Stevenson, Knott, Overlander, & McDonald 2000)
• Signaling an Explanation coherence relation (Hobbs 1979; Kehler 2002)

3. Using coherence to mode next-mention biases
We generalize Rohde, Kehler, & Elman’s (2007) pronoun model to next mention: Biases towards upcoming coherence relations (CRs) combine with biases for which event participant will be mentioned again, conditioned on coherence

(3) P(next_mention = referent) = \sum_{CR} P(CR) \cdot P(next_mention | referent) \cdot CR

Because
P(CR=Explanation) = 1 with ‘because’, but P(CR=Explanation) < 1 in full-stop
Next-mention bias, P(next_mention | Explanation), is predicted to remain constant across Explanations – with both ‘because’ and full stop Explanations

4. Story continuation experiment
2 x 3 design: verb type (IC vs. Non-IC) x continuation type (full stop vs. because vs. dialog prompt – dialog results not discussed here)

Materials: 40 IC verbs (20 IC-1, 20 IC-2) and 40 Non-IC verbs
Evaluation: judges annotated for next mention & coherence relation

Results: By categorizing responses by coherence relation, we localize the previously reported IC bias to Explanation relations. We find an additional IC bias concerning P(Explanation). This bias has gone unnoticed because previous work has not categorized responses by coherence.

6. IC-1 Results
Next-mention biases were statistically indistinguishable when only ‘because’ prompts and freely generated Explanations were considered (F(1,70)=0.0221, p=0.8822; F(1,19)=0.032, p=0.8606)

Prompt: ‘because’
P(next_mention = NP1 | ‘because’) = p(next_mention=NP1 | Explanation)

6. IC-2 Results
Again, next-mention biases statistically indistinguishable when only Explanations are considered ‘because’ or freely generated (F(1,61)=1, p=0.982; F(1,36)=1.4598, p=0.2348)

Prompt: full stop

7. IC-2 Results
Additionally, responses with full stop prompt (Caramazza, Grober, Garvey, Yates 1974; McKoon, Greene, Ratcliff 1993)

Because
P(next_mention = NP1 | ‘because’) = p(next_mention=NP1 | Explanation)

8. Non-IC Results

9. A new IC bias
IC verbs create an expectation regarding the direction the discourse is likely to take – specifically a bias towards an upcoming Explanation

Findings for full-stop prompt: IC verbs yield more Explanation continuations than do Non-IC verbs

10. Conclusions
Like Rohde et al.’s results, overall statistics conceal a consistent system of stronger biases once coherence relations are conditioned on.

In contrast to previous results:
• Connective alone does not affect referent salience – mediated by coherence
• There are actually two strong biases that differentiate IC and Non-IC verbs:
P(CR=Explanation) is high for IC-1 and IC-2
P(next_mention = NP1 | Explanation) is high for IC-1 and low for IC-2
The presence of a second bias had gone unnoticed because previous studies had not categorized their data by coherence.

References

Contact: hannah@ling.ucsd.edu