DETEC2019, September 27-28, 2019, Berlin

Pronominalization and expectations for re-mention
Evidence from benefactives

Jet Hoek,[1] Andrew Kehler,[2] & Hannah Rohde[3]
[1]University of Cologne, [2]UCSD, [3]The University of Edinburgh
Contact: jhoek@uni-koeln.de

LEVERHULME TRUST

Abstract

Coreference provides a window into people's inferences and expectations about relationships that hold across sentences. Different approaches to coreference place different emphasis on the roles of meaning (Winograd 1972; Hobbs 1979) and form (Grosz et al. 1995)—two components which are combined in the Bayesian Model put forward by Kehler et al. (2008). The Bayesian Model, in its strong form, posits the independence of a referent's predictability for re-mention and its likelihood of being mentioned with a pronoun. However, evidence regarding this independence is mixed.

We use a **new context type with three referents** to test:

(i) whether predictability influences pronominalization

(ii) whether Bayes' Rule captures the relationship between pronoun interpretation and production

In a story continuation experiment, we replicate two previously-established findings: that the presence of a pronoun boosts the proportion of continuations about the subject, and that grammatical role influences pronominalization. We find no evidence of any dependence between predictability and pronominalization.

While both these findings are in line with the Bayesian model, the Bayesian model is not the best quantitative fit for our data when evaluated against two other models. Follow-up studies will evaluate if this is due to the 3-referent context or to the specific construction we tested.

1. Goal

To replicate/extend findings on pronoun production and pronoun comprehension to a new type of sentence frame with more than two referents.

2. Models of coreference

Mirror Model
Listeners base their interpretation decisions on their estimates of the speaker's likelihood to use a pronoun to mention particular referents
(Ariel 1990; Gundel et al. 1993)

p(referent | pronoun) ~ p(pronoun | referent)

Expectancy Model
Listeners' expectations about who will be mentioned next determines their interpretation of a subsequent pronoun (Arnold 2001)

p(referent | pronoun) ~ p(referent)

Bayesian Model
Incorporates both of these components: an expectation about which referent will be re-mentioned (the prior) and an estimate of how likely a speaker is to use a pronoun when re-mentioning a particular referent (the likelihood)
(Kehler et al. 2008; Kehler & Rohde 2013; Rohde & Kehler 2014)

$p(\text{referent} \mid \text{pronoun})_{\text{INTERPRETATION}} \sim p(\text{referent})_{\text{PRIOR}} * p(\text{pronoun} \mid \text{referent})_{\text{LIKELIHOOD}}$

3. Strong versus weak Bayes

(1) a. John scolded Bob. He _____ [pronoun prompt condition]
b. John scolded Bob. _____ [free prompt condition]

In story continuations involving items like (1), a pronoun is more likely to be interpreted as referring to the one referent (Bob in 1a), but more likely to be produced to refer to the other (John in 1b) (Stevenson et al. 1994)

<u>The Bayesian model captures this asymmetry</u>

In its **strong form**, the Bayesian model separates the discourse features that influence the prior and the likelihood:

meaning drives the *prior*
topicality drives the *likelihood*

→ This complete separation has been contested in recent work that shows that the likelihood of pronominalization increases for referents with a higher prior (e.g., Rosa & Arnold 2017)

In its **weak form**, the Bayesian model states that the **pronoun production and interpretation are related by Bayesian principles.**

4. Research questions

Most of the research on pronoun production / interpretation has focused on sentence frames with two referents. Results appear to differ between implicit causality verbs (Rohde 2008; Fukumura & van Gompel 2010; Kehler & Rohde 2014) and studies with transfer-of-possession verbs (Rosa & Arnold 2017)

In a new context type with three referents, we test:
(i) whether predictability influences pronominalization (strong vs weak Bayes)
(ii) whether Bayes' Rule captures the relationship between pronoun interpretation and production

5. Story continuation experiment

Task: Write a natural continuation for the experimental prompts

Materials: 30 prompts containing an agent, patient, and a benefactive:

(3) a. Adam scolded Diana for Russell. He _____ [pronoun prompt condition]
b. Adam scolded Diana for Russell. _____ [free prompt condition]

We counter-balanced which potential referents were gender-matched (NP1&NP2, NP1&NP3, NP2&NP3)

Coding:
- Who the continuation is about [both conditions]
- What form of referring expression is used [free prompt condition only]

Number of participants: 83

6. (i) Does predictability influence pronominalization?

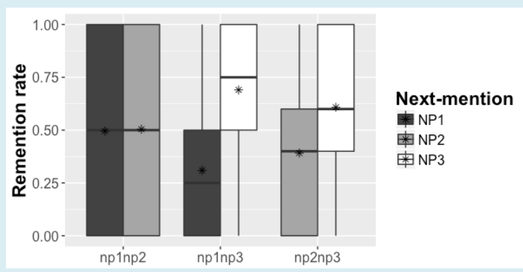

Figure 1. Re-mention rates in the free prompt condition, per ambiguous pair; asterisks show participant means.

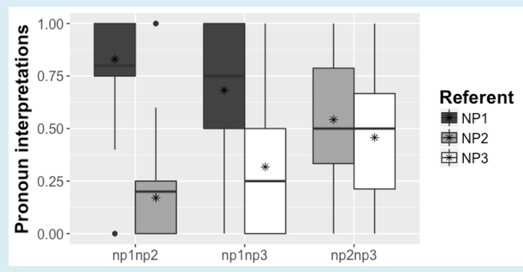

Figure 2. Pronoun interpretation in the pronoun prompt condition, per ambiguous pair.

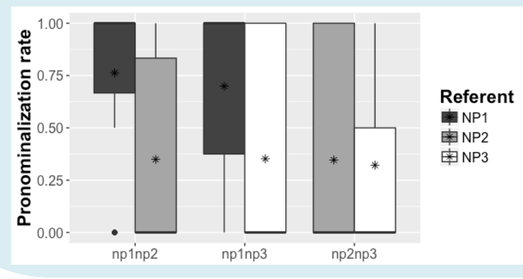

Figure 3. Pronoun rates by referent in free prompt condition, per ambiguous pair.

<u>We replicate two well-established findings:</u>

The pronoun prompt yields more NP1 continuations (β=1.52, p<.001; compare Figs 1&2)

The subject is preferentially rementioned by means of a pronoun (See Fig 3)

<u>For question (i), compare Figs 1 and 3:</u>

Remention rates of NP1 and NP2 do not differ (β=0.22, p=.53)**, but their pronominalization rates do**; (β=-3.26, p<.001)

Remention rates of NP2 and NP3 differ (β=1.12, p<.001)**, but their pronominalization rates do not**; (β=-0.19, p=.42)

We thus find no evidence of any dependence between predictability and pronominalization

7. (ii) Does Bayes' Rule rule?

Following Rohde & Kehler (2014), we used the full-stop continuations to calculate Bayes-derived estimates of *p(referent | pronoun)* via the prior *p(referent)* and likelihood *p(pronoun | referent)*, as well as estimates for the Expectancy Model (prior) and the Mirror Model (normalized likelihood).

Does Bayes' Rule rule?

→ In this context, not really

	NP1NP2		NP1NP3		NP2NP3	
	NP1	NP2	NP1	NP3	NP2	NP3
Observed	**0.82**	**0.18**	**0.73**	**0.27**	**0.57**	**0.43**
Bayes	0.72	0.28	0.57	0.43	0.40	0.60
Expectancy	0.46	0.54	0.34	0.66	0.41	0.59
Mirror	0.77	0.23	0.70	0.30	0.42	0.58

Table 1. Mean observed pronoun interpretation rate and estimates for all three models calculated **over items**, per referent, per ambiguous pair.

	items	participants
Bayes	R^2=.122, p<.0001	**R^2=.084, p<.0001**
Expectancy	R^2=.003, p=.039	R^2=.021, p=.509
Mirror	**R^2=.377, p<.0001**	R^2=.075, p<.0001

Table 2. Correlations between each model's predicted pronoun interpretation and the observed pronoun interpretation, calculated over items and over participants.

8. Discussion

As in earlier work, the Bayesian Model's correlation with the observed pronoun interpretation is higher than that of the Expectancy Model. In contrast to earlier work, however, the Mirror Model provided the best fit to the observed data, *at least in the by-items analysis.*

→ Is the Bayesian model's poor performance due to the presence of three referents, or does it have something to do with the specific sentence frame (Benefactives)?

Follow-up 1: Benefactive sentence frames with 2 human referents (n=85)

(4) **Adam** scolded the clean-up crew for **Russel**.

<u>Preliminary results</u>:

	NP1 (subj)	NP2 (ben)
Observed	**0.68**	**0.32**
Bayes	0.64	0.36
Expectancy	0.23	0.77
Mirror	0.87	0.13

Table 3. Mean observed pronoun interpretation rate and estimates for all three models in the follow-up experiment calculated over items, per referent.

	items	participants
Bayes	**R^2=.719, p<.0001**	**R^2=.348, p<.0001**
Expectancy	R^2=.311, p<.0001	R^2=.008, p=.334
Mirror	R^2=.714, p<.0001	R^2=.282, p<.0001

Table 4. Correlations between each model's predicted pronoun interpretation and the observed pronoun interpretation, calculated over items and over participants.

→ Bayes is back!
But what does this mean?

References
Arnold (2001). *Discourse Processes*,137-162. **Ariel (1990).** *Accessing noun phrase antecedents*. London: Routledge. **Fukumura & van Gompel (2010).** *Jrnl of Memory and Language*, 52-66. **Grosz et al. (1995).** *Computational Linguistics* 21, 203–225. **Gundel et al. (1993)**. *Language*, 274–307. **Hobbs (1979).** *Cognitive Science*, 67–90. **Kehler et al. (2008).** *Jrnl of Semantics*, 1-44. **Kehler & Rohde (2013)**. *Theoretical Linguistics*, 1-37. **Rohde (2008**). *PhD Thesis*. **Rohde & Kehler (2014)**. *Language, Cognition, and Neuroscience*, 1-16. **Rosa & Arnold (2017)**. *Jrnl of Memory and Language*, 43-60. **Stevenson et al. (1994).** *Language and Cognitive Processes*, 519–548. **Winograd (1972).** *Natural language understanding*. NY: Academic Press.

Acknowledgments This work was supported by a Leverhulme Trust Prize in Languages & Literatures.