

EVALUATING BAYESIAN MODELS OF PRONOUN INTERPRETATION WITH MANDARIN STORY CONTINUATION DATA

Conghui Zhu & Fuyun Wu (Shanghai Jiao Tong University)

The Bayesian account of pronoun interpretation posits two determinants for pronoun interpretation biases ^{[1][2]}: (a) $P(\text{pronoun}|\text{referent})$, the probability that people use a pronoun to refer to an entity, and (b) $P(\text{referent})$, the probability that an entity will be mentioned again. This account comes with two competing models: the Expectancy Model where pronoun interpretation is affected only by $P(\text{referent})$, and the Mirror Model where pronoun interpretation is affected only by $P(\text{pronoun}|\text{referent})$ ^[1]. In a strong version of Bayesian Models ^[3], factors conditioning $P(\text{referent})$ are prescribed as primarily semantic and pragmatic, and factors conditioning $P(\text{pronoun}|\text{referent})$ are grammatical and/or information structural. But this assertion is odd, since information structure is (at least partly) a pragmatic phenomenon ^[4]. To evaluate different models of pronoun interpretation, we conducted two Mandarin story continuation experiments, examining how the two information structure roles, specifically Topic and Focus, affect pronoun production and interpretation.

We created 24 sets of stories with rich contextualisation, each totalling 6 clauses (see ex.(1) in English translations). Two referents, NP1 and NP2, were introduced in succession into each story. We manipulated the information status of NP1 by placing it before (Topic) or after (nonTopic) the conjunction 'if', and that of NP2 by asking a question (Focus) or adding a statement (nonFocus) before NP2. **EXP 1** (N=40) was free-prompt: participants could freely choose referring expressions for either NP1 or NP2. Hence, we could compute production probabilities. **EXP 2** (N=40) was pronoun-prompted: participants already interpreted the given pronoun, thus their continuation reflected interpretation biases.

We built generalised linear mixed effect models on $P(\text{referent})$ for both experiments, and on $P(\text{pronoun}|\text{referent})$ for EXP 1. Both experiments yielded more references to NP1 when it was topicalized than not (EXP 1: $\beta = 0.60$, $SE = 0.26$, $t = 2.33$; EXP 2: $\beta = 0.66$, $SE = 0.22$, $t = 2.98$). EXP 2 also showed more references to NP2 when it was focused than not ($\beta = -0.57$, $SE = 0.22$, $t = -2.57$). No interaction was found in either experiment. Clearly, topicalization can affect $P(\text{referent})$, contrary to previous findings ^[5]. This discrepancy may be attributed to our rich story context, which is reportedly crucial for effects of predictability on pronoun use ^[6].

Of the three models, Bayesian makes the best prediction on the pronoun interpretation bias (R^2 of model prediction and observed data: Bayesian: 0.40, Expectancy: 0.31, Mirror: 0.16). Note that EXP 1 showed no effect of either topic or focus on $P(\text{pronoun}|\text{referent})$, which corresponds to the part of Bayesian Model that differs the Expectancy Model. Yet the Bayesian Model outperforms the Expectancy Model, suggesting that the reportedly lack of interaction between $P(\text{referent})$ and $P(\text{pronoun}|\text{referent})$ ^[5] does not necessarily imply their disassociation, but an information structural factor (topicality) might affect both probabilities.

Taken together, our Mandarin story continuation data demonstrated effects of Topic and Focus on pronoun resolution and the effect of topic on the $P(\text{referent})$. Our model comparison results stand against the strong version of the Bayesian Model where information structure only affects $P(\text{pronoun}|\text{referent})$, but support a weak form of the Bayesian account such that pronoun interpretation and production follow Bayesian principles ^[1].

(1) Lead-in clause (Clause 1): "Comptroller Lin-Xue (NP1) went to an insurance company to audit at noon."

Focus Manipulation (Clause 2-4):

Focus: "Who catered reception?"

"Assistant Zhang-Chao (NP2, Focus) catered reception, he is familiar with the processes."

Non-focus: "The accounting work should be finished by next Monday."

"Assistant Zhang-Chao (NP2, nonFocus) catered reception, he is familiar with the processes."

Topic Manipulation (Clause 5):

Topic: "Lin-Xue (topic), if noticed the gaps and omissions of accounting items,"

Non-topic: "If Lin-Xue (non-topic) noticed the gaps and omissions of accounting items,"

Final to-be-completed clause (Clause 6): "then _____." (**EXP 1**) or "then *ta* ('he'/'she') _____." (**EXP 2**)"

References: [1] Kehler, Kertz, Rohde & Elman. (2008). *J Semant*, 25(1), 1-44. [2] Kehler & Rohde (2013). *Theor Linguist*, 39(1-2), 1-37. [3] Kehler & Rohde. (2019). *J. Pragm*, 154, 63-78. [4] Gundel. (2012). In Allan & Jaszczolt (Eds.), *The Cambridge Handbook of Pragmatics*. 585-598. [5] Rohde & Kehler. (2014). *LCN*, 29(8), 912-927. [6] Demberg, Kravtchenko, & Loy. (2023). *JML*, 130, 104413.