Does the interpretation of null pronouns in Mandarin follow a Bayesian approach?

Suet Ying Lam & Heeju Hwang

Department of Linguistics
The University of Hong Kong

34th CUNY Human Sentence Processing Conference

March 5, 2021

Models in pronoun interpretation

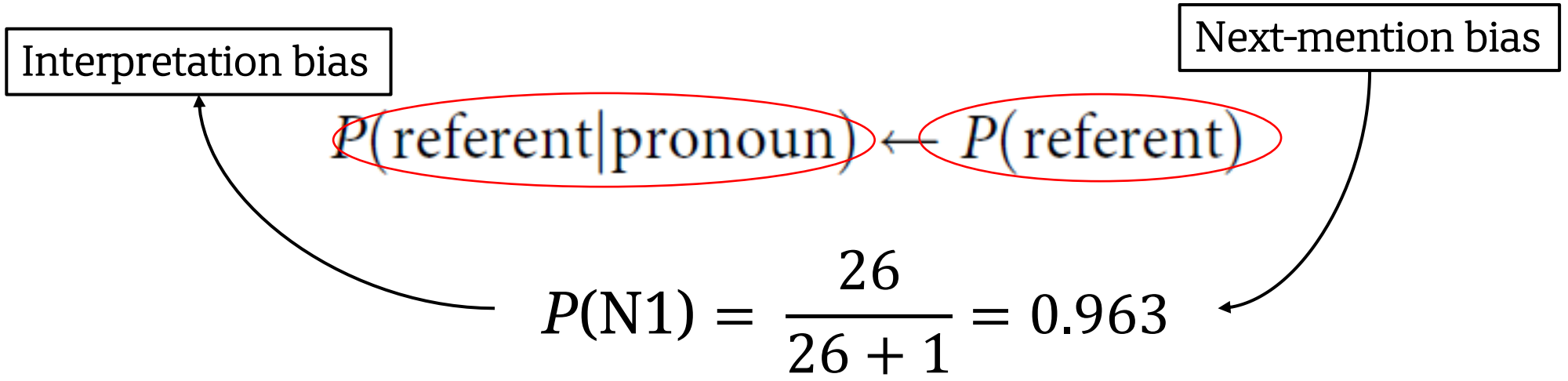
Expectancy Hypothesis
(e.g., Arnold, 2001)

Mirror Model
(e.g., Ariel, 1990; Givón, 1983; Gundel et al., 1993)

Bayesian Approach
(e.g., Kehler et al., 2008)

Models in pronoun interpretation

1. Expectancy Hypothesis



小玲吓到了嘉怡，因为…
Xiaoling frightened Jiayi, because…

Continuations about Xiaoling (N1): 26
Continuations about Jiayi (N2): 1

Models in pronoun interpretation

Expectancy Hypothesis
(e.g., Arnold, 2001)

Mirror Model
(e.g., Ariel, 1990; Givón, 1983; Gundel et al., 1993)

Bayesian Approach
(e.g., Kehler et al., 2008)

Models in pronoun interpretation

2. Mirror Model

Interpretation bias

Production bias

$$P(\text{referent}|\text{pronoun}) \leftarrow \frac{P(\text{pronoun}|\text{referent})}{\sum_{\text{referent}\in\text{referents}} P(\text{pronoun}|\text{referent})}$$

$$\frac{P(\text{overt}\,|\,\text{N1})}{P(\text{overt}\,|\,\text{N1}) + P(\text{overt}\,|\,\text{N2})} = \frac{0.615}{0.615 + 0}$$

小玲吓到了嘉怡，因为…
Xiaoling frightened Jiayi, because…

Continuations about Xiaoling (N1): 26
Using Overt pronoun: 16

$$P(\text{overt}\,|\,\text{N1}) = \frac{16}{26} = 0.615$$

Continuations about Jiayi (N2): 1
Using Overt pronoun: 0

$$P(\text{overt}\,|\,\text{N2}) = \frac{0}{1} = 0$$

Models in pronoun interpretation

Expectancy Hypothesis
(e.g., Arnold, 2001)

Mirror Model
(e.g., Ariel, 1990; Givón, 1983; Gundel et al., 1993)

Bayesian Approach
(e.g., Kehler et al., 2008)

Models in pronoun interpretation

3. Bayesian approach

Interpretation bias

Next-mention bias
Production bias

$$P(\text{referent}|\text{pronoun}) = \frac{P(\text{pronoun}|\text{referent})P(\text{referent})}{\sum_{\text{referent}\in\text{possible referents}} P(\text{pronoun}|\text{referent})P(\text{referent})}$$

$$\text{P(N1|overt)} = \frac{(0.963 \times 0.615)}{(0.963 \times 0.615)+(0\times 0.037)}$$

小玲吓到了嘉怡，因为…
Xiaoling frightened Jiayi, because…

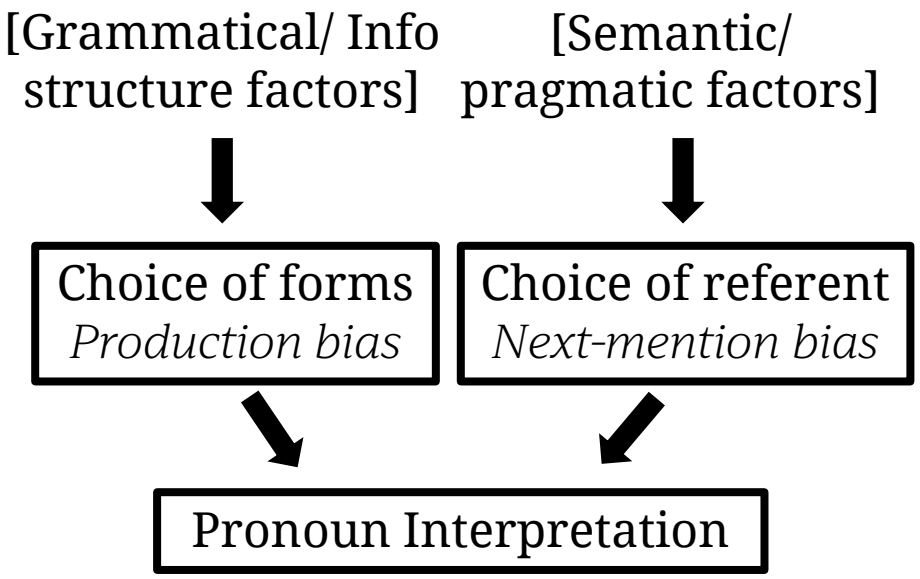
$$P(\text{N1}) = \frac{26}{27} = 0.963 \quad P(\text{N2}) = \frac{1}{27} = 0.037$$

$$P(\text{overt}|\text{N1}) = \frac{16}{26} = 0.615 \quad P(\text{overt}|\text{N2}) = \frac{0}{1} = 0$$

Models in Pronoun Interpretation

Bayesian Model

Asymmetry between interpretation and production/next-mention bias



Performed the best in:

- ✓ English overt pronouns (Rhode & Kehler, 2014; Kehler & Rhode, 2019)
- ✓ Chinese overt pronouns (Zhan et al., 2020)
- ✓ German personal pronouns (Patterson et al., 2020)

Research question:
Which model best explains the interpretation of null pronoun?

Overt and null pronouns are different:

- Null pronoun interpretation is more strongly subject-biased (e.g., Japanese: Ueno & Kehler, 2016; Mandarin: Zhang, 2018)
- Null pronoun interpretation is not sensitive to pragmatic factors like aspect in Japanese (Ueno & Kehler, 2016)

Methods

- Two passage-continuation experiments on overt pronouns (exp.1) and null pronouns (exp.2)

Implicit causality (IC) verbs: Verb Bias x Prompt Type, with connective "because"

Subject (N1)-biased verb	Object (N2) -biased verb	→ Next-mention bias
小玲吓到了嘉怡，因为…	嘉怡害怕小玲，因为…	
Xiaoling frightened Jiayi, because	Jiayi scares Xiaoling, because	

Transfer-of-possession (TOP) verbs: Verb Bias x Prompt Type, with connective "so"

Subject (N1)-biased verb	Object (N2)-biased verb
立强从小刚那里收到了一个包裹，所以…	小刚向立强寄了一个包裹，所以…
Liqiang received a package from Xiaogang, so	Xiaogang sent a package to Liqiang, so

Methods

- Two passage-continuation experiments on overt pronouns (exp.1) and null pronouns (exp.2)

Implicit causality (IC) verbs: Verb Bias x Prompt Type, with connective "because"

Subject (N1)-biased verb

小玲吓到了嘉怡，因为…
Xiaoling frightened Jiayi, because

Object (N2) -biased verb

嘉怡害怕小玲，因为…
Jiayi scares Xiaoling, because

→ Next-mention bias

Free prompt

因为 (free)…
because…

Pronoun prompt

因为她 (overt)…
because she…

因为想 (null)…
because wants to/think…

Free: Next-mention and production bias
Pronoun: Interpretation bias

Transfer-of-possession (TOP) verbs: Verb Bias x Prompt Type, with connective "so"

Subject (N1)-biased verb

立强从小刚那里收到了一个包裹，所以…
Liqiang received a package from Xiaogang, so

Object (N2)-biased verb

小刚向立强寄了一个包裹，所以…
Xiaogang sent a package to Liqiang, so

Free prompt

所以 (free)…
so…

Pronoun prompt

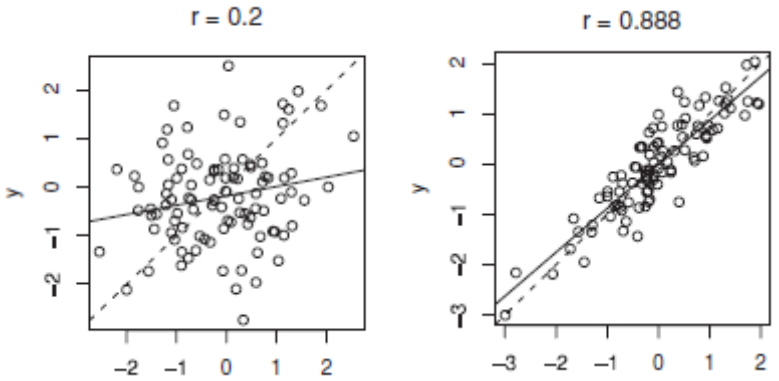
所以她 (overt)…
so she…

所以想 (null)…
so wants to/think…

Evaluation Metrics

Correlation: R-squared (R2)

- How well the observed values are replicated by the predicted value



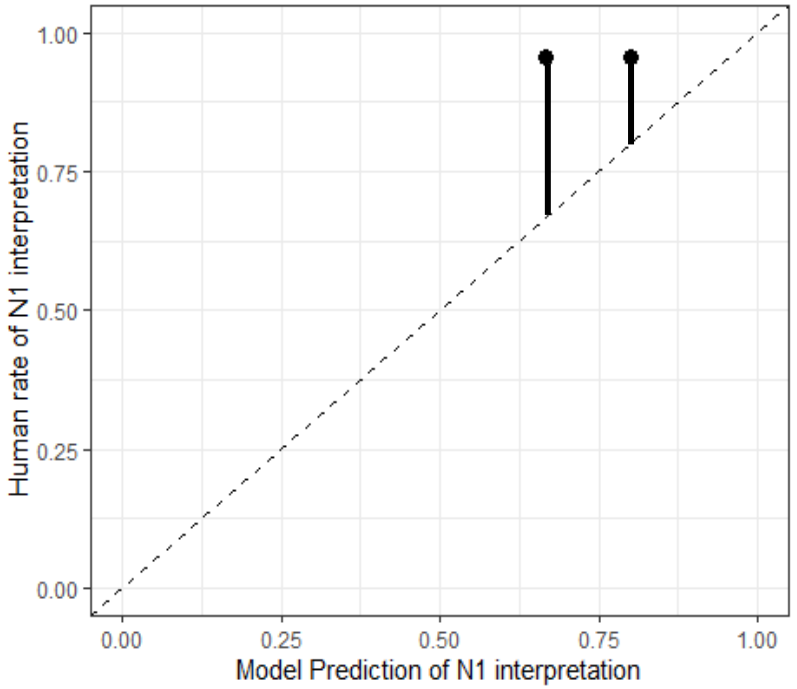
(Baayen, 2008; P.88)

Prediction error: Mean squared error (MSE)

- Reflects how far away the prediction is from the observation

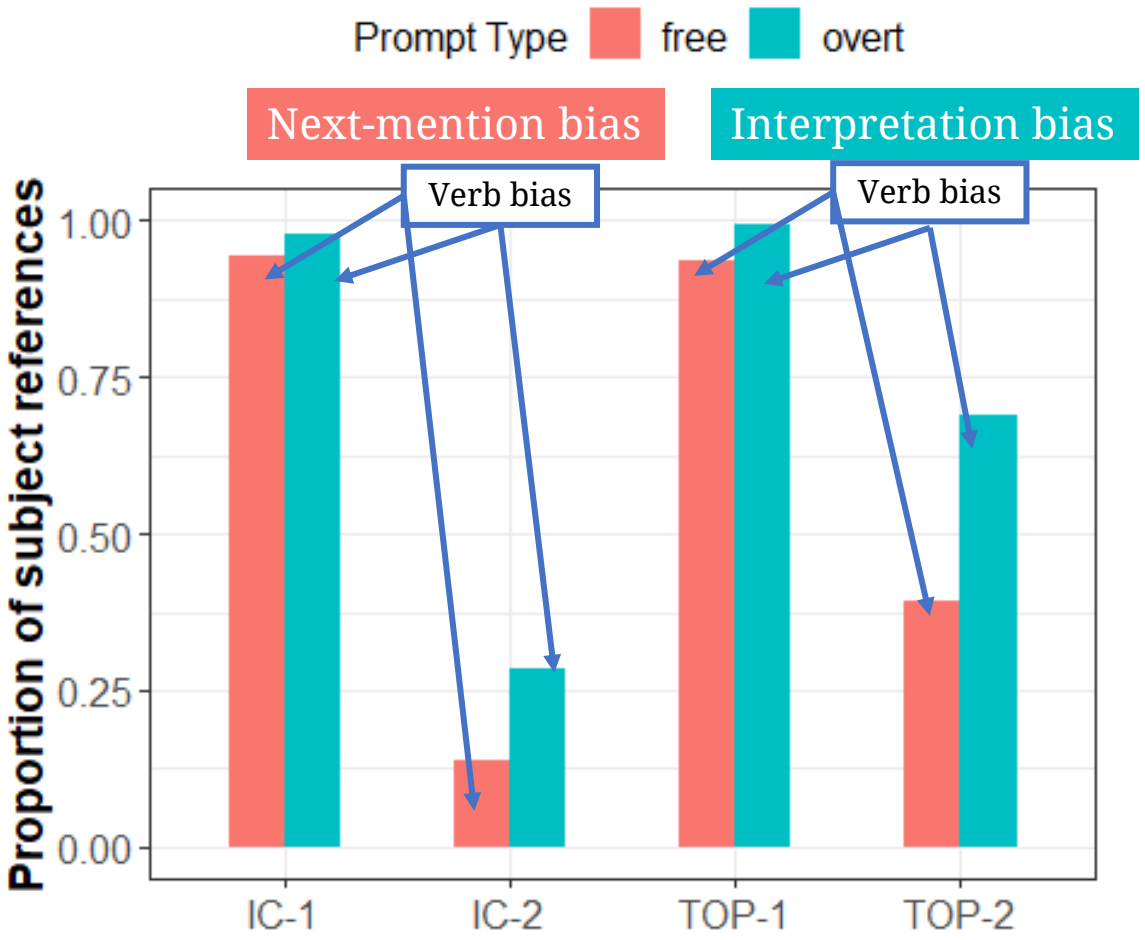
Prediction error: Average Cross Entropy (ACE)

- Reflects how far away the prediction is from the observation in a logarithmic scale
- Gives more weight to extreme deviations



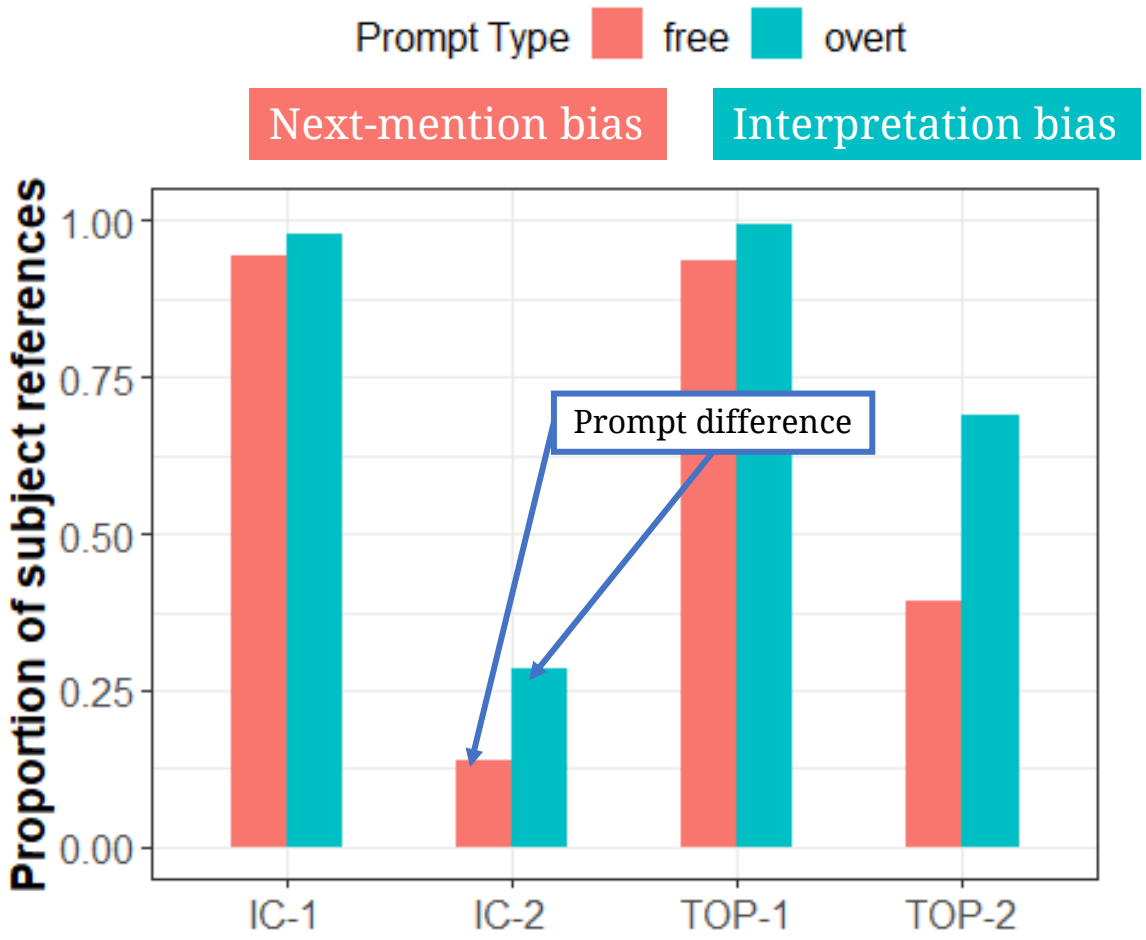
Overt pronoun (Experiment 1) results: Choice of referents and forms

Choice of referent



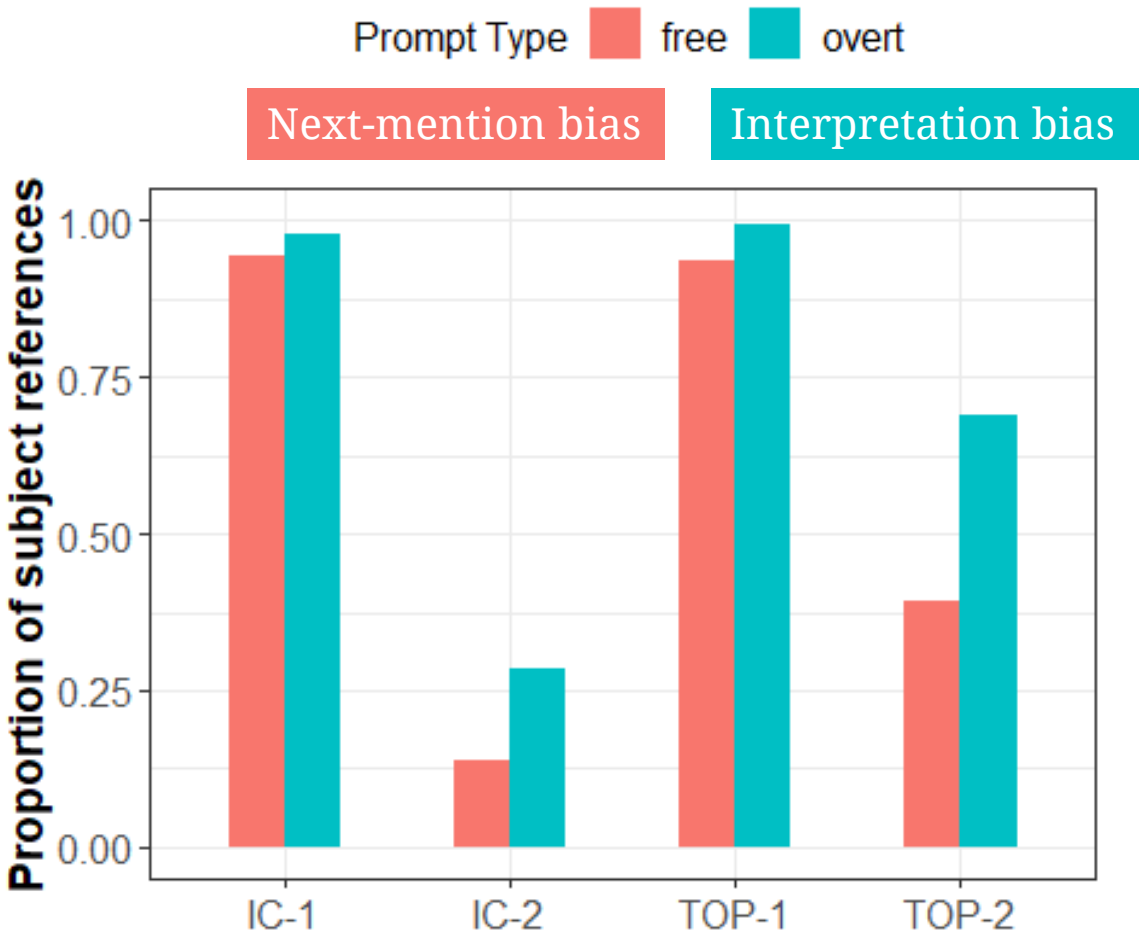
Overt pronoun (Experiment 1) results: Choice of referents and forms

Choice of referent

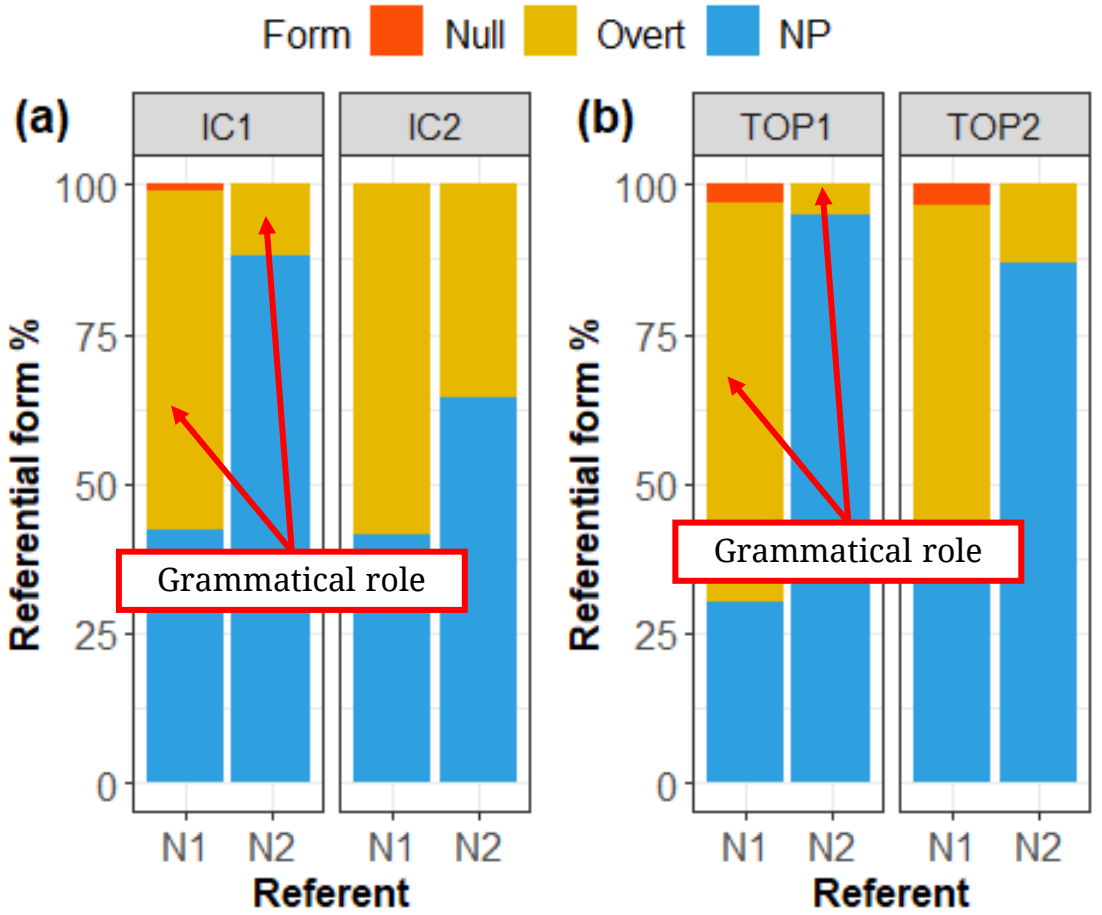


Overt pronoun (Experiment 1) results: Choice of referents and forms

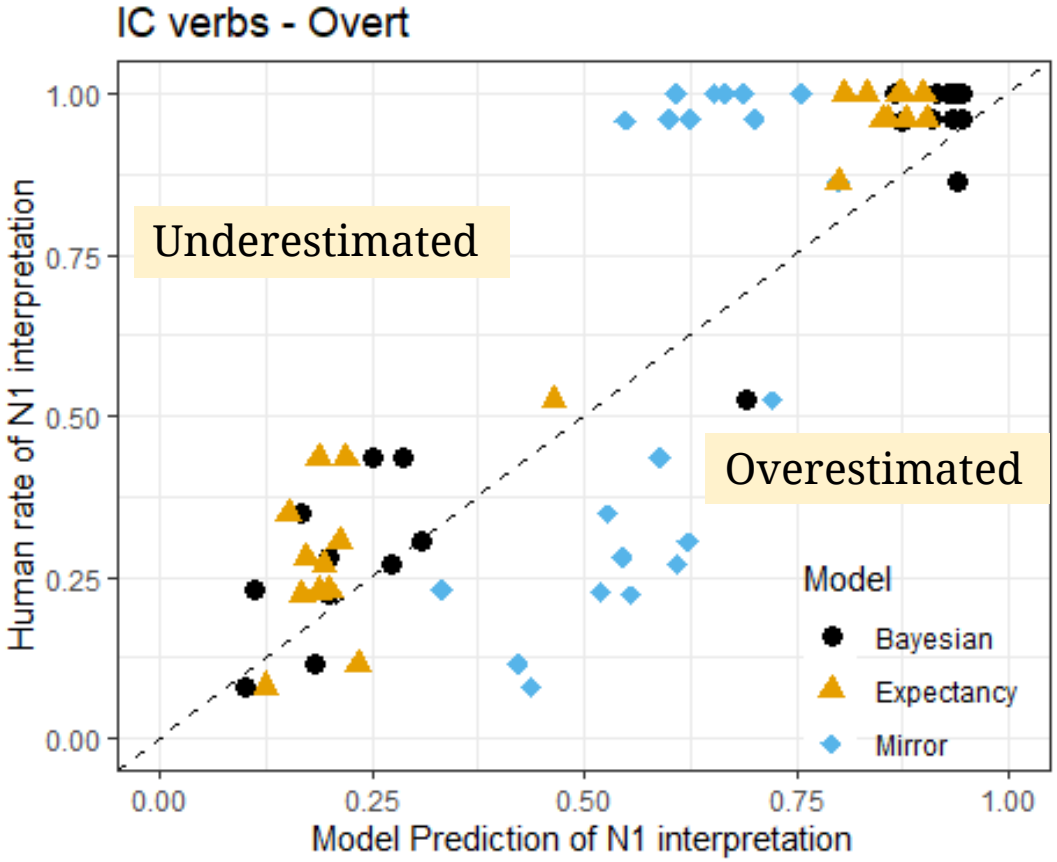
Choice of referent

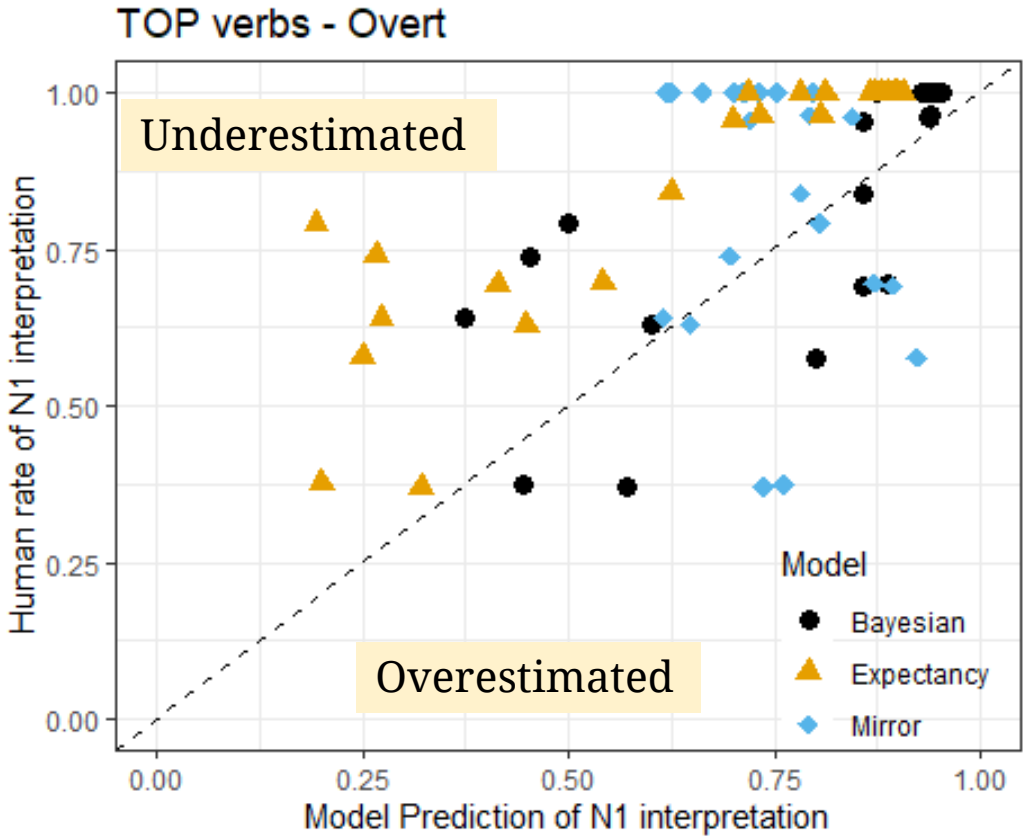


Choice of referential form (Production bias)



Model Evaluation – Overt pronoun





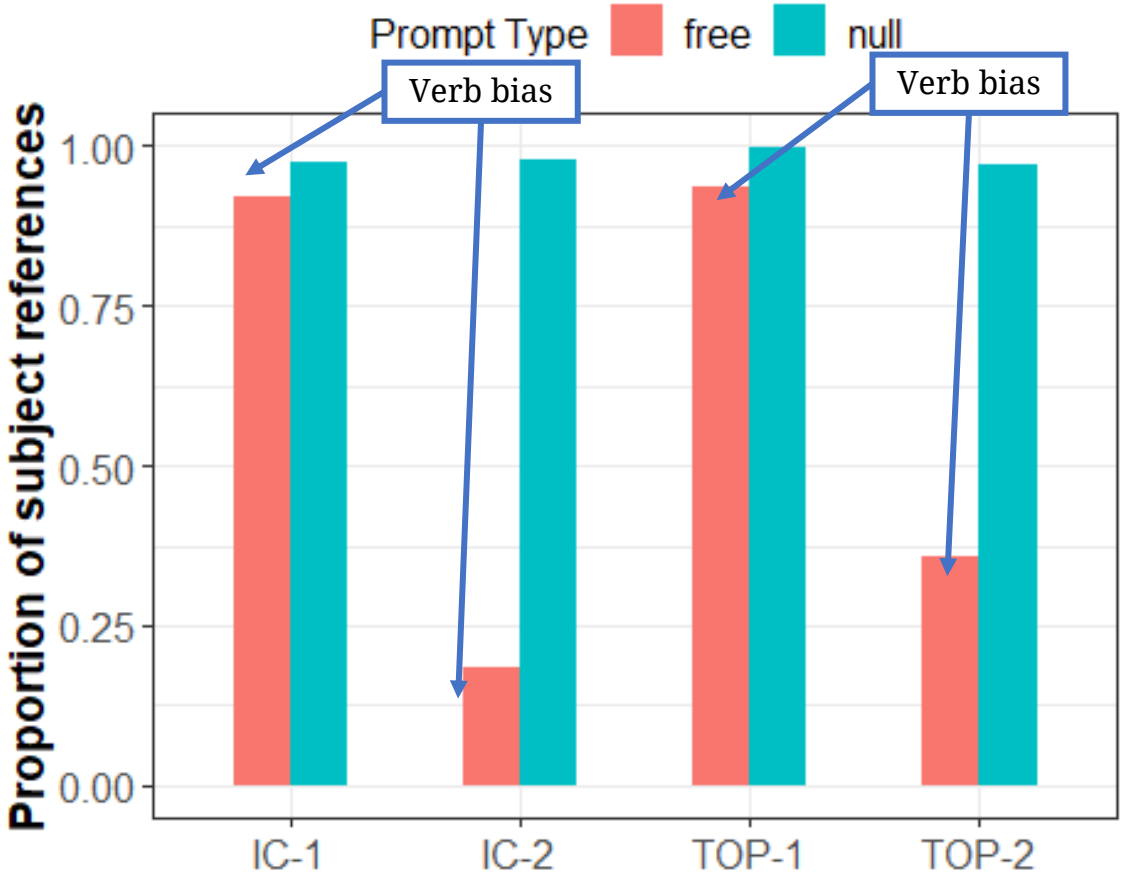
IC Overt	**Bayesian**	**Expectancy**	**Mirror**
R^2	0.950***	**0.952*** **	0.491***
MSE	**0.009**	0.016	0.085
ACE	0.804	0.888	**0.521**

TOP Overt	**Bayesian**	**Expectancy**	**Mirror**
R^2	0.585***	**0.772*** **	0.080
MSE	**0.019**	0.060	0.068
ACE	**0.253**	0.575	0.309

***: $p < .001$; **: $p < .01$; *: $p < .05$;

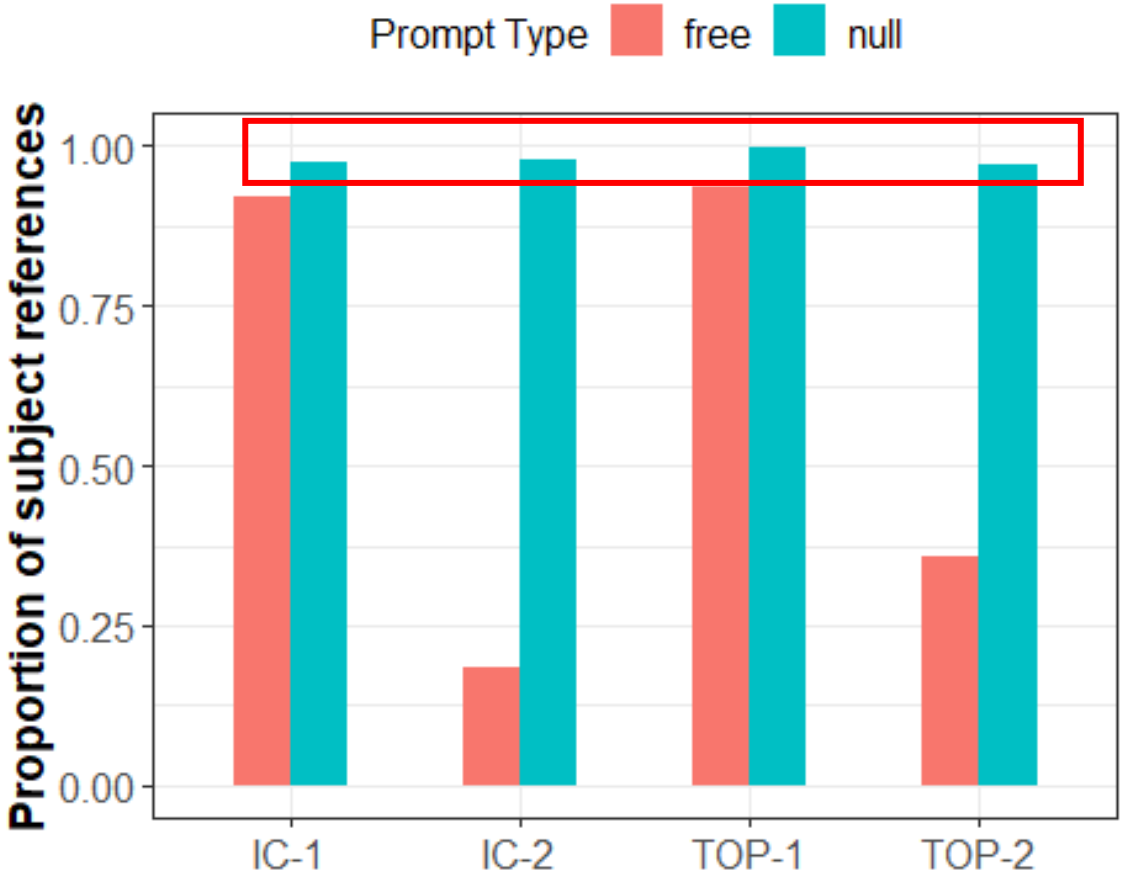
Results: Experiment 2 (Null pronoun) - Choice of referents and forms

Choice of referent
(Next-mention/Interpretation bias)



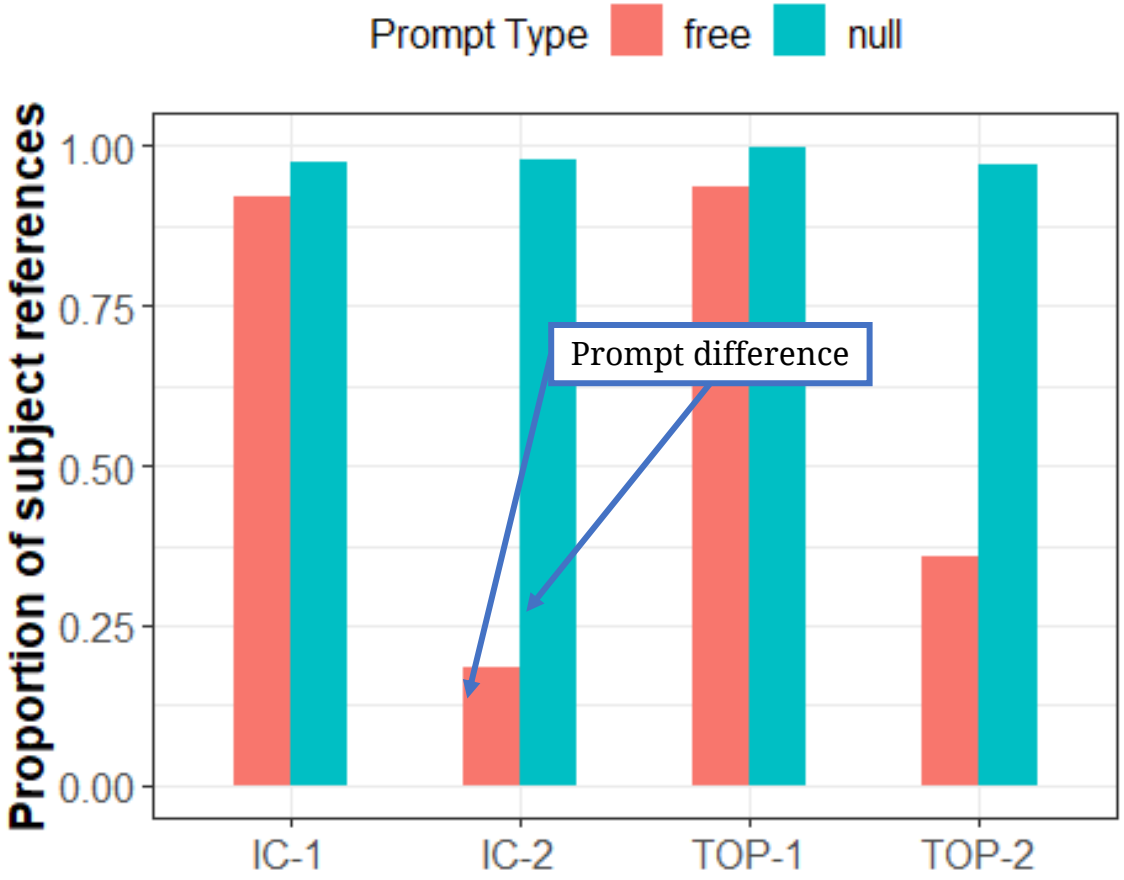
Results: Experiment 2 (Null pronoun) - Choice of referents and forms

Choice of referent
(Next-mention/Interpretation bias)

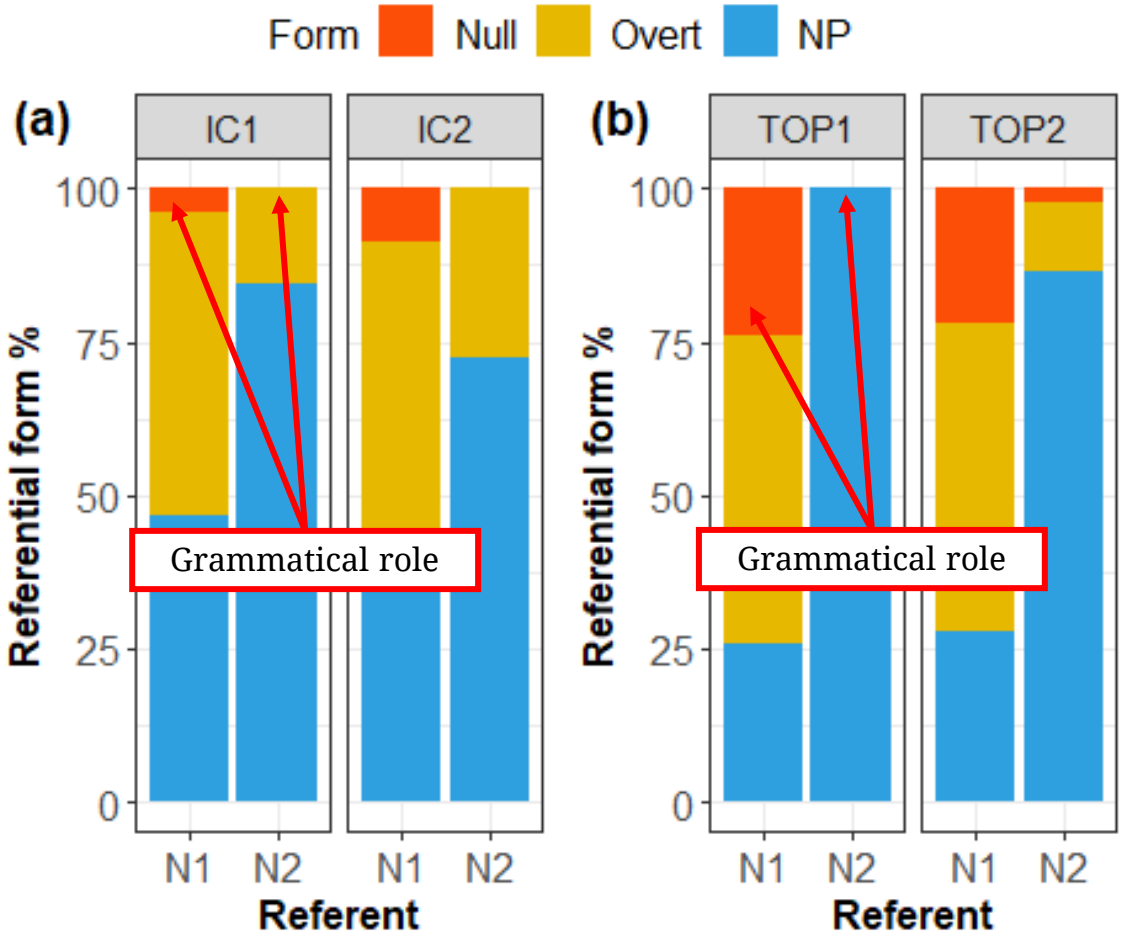


Results: Experiment 2 (Null pronoun) - Choice of referents and forms

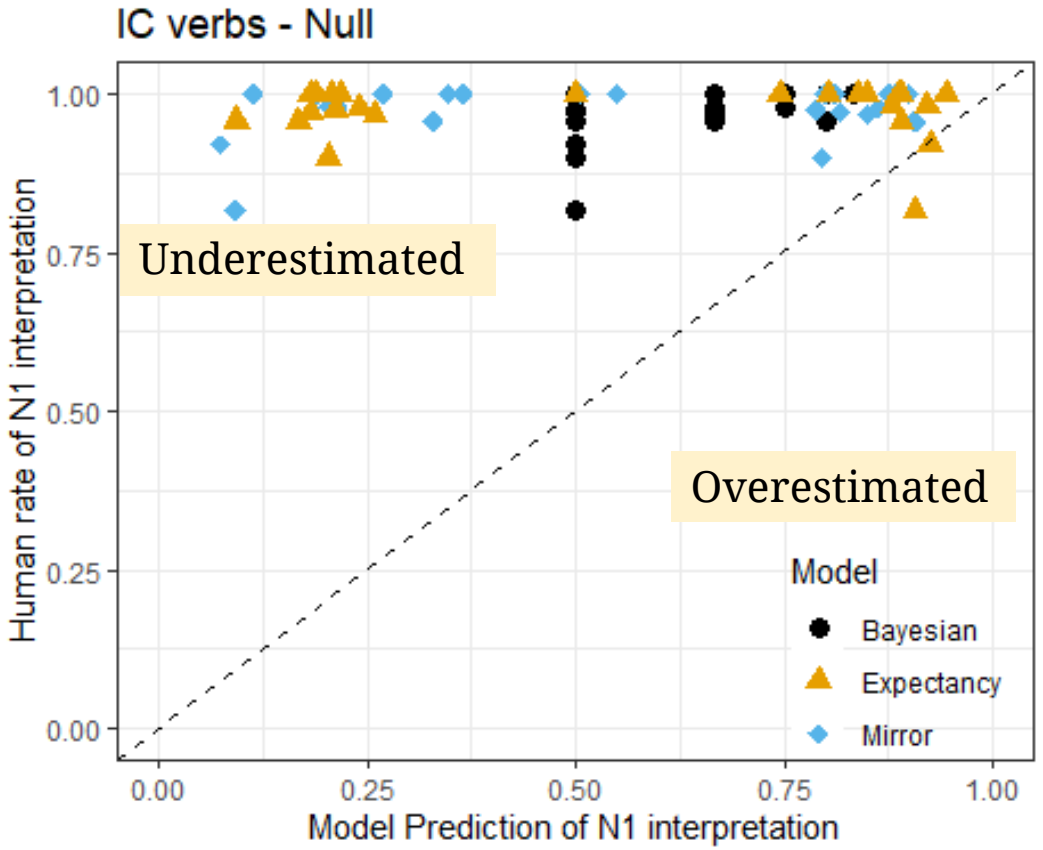
Choice of referent (Next-mention/Interpretation bias)

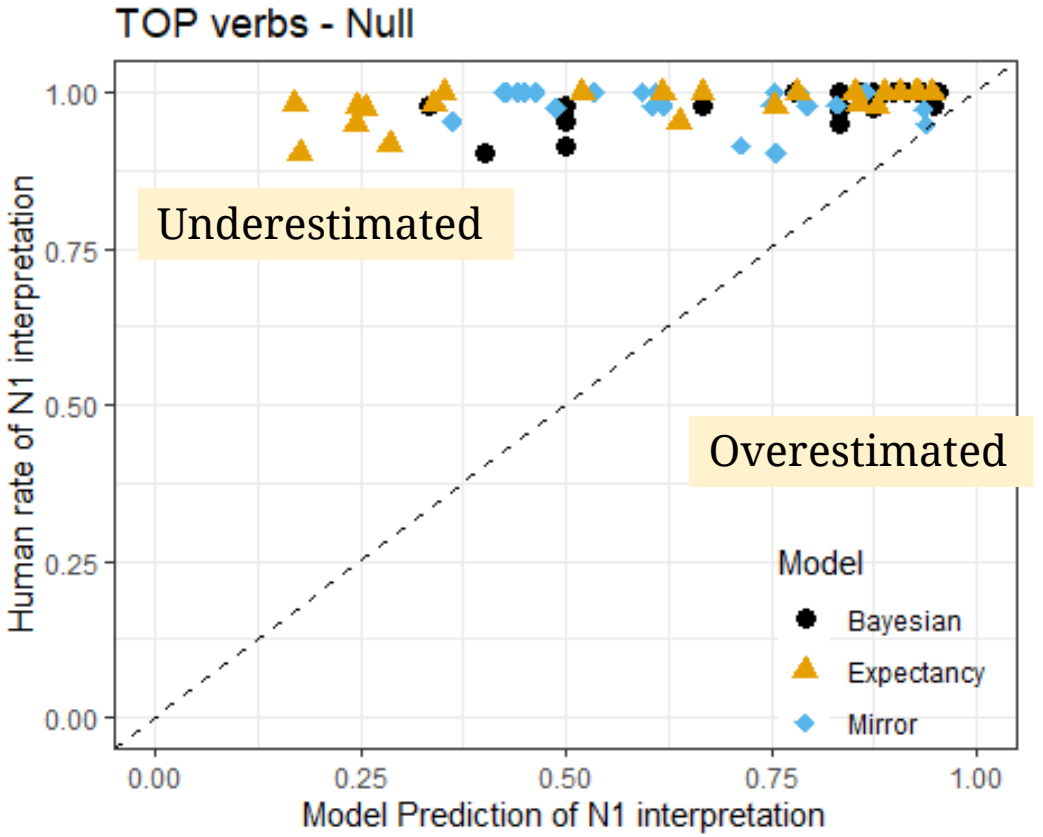


Choice of referential form (Production bias)



Model Evaluation – Null pronoun





IC Null	**Bayesian**	**Expectancy**	**Mirror**
R^2	**0.253***	0.004	0.045
MSE	**0.114**	0.297	0.262
ACE	**0.445**	0.858	0.820

TOP Null	**Bayesian**	**Expectancy**	**Mirror**
R^2	**0.488*****	0.383**	0.041
MSE	**0.067**	0.197	0.137
ACE	**0.277**	0.606	0.452

***: $p < .001$; **: $p < .01$; *: $p < .05$;

Conclusion: Take-home messages

Overt pronoun

Well-explained by the Bayesian approach

- Except that verb bias also affects the production of overt pronoun in TOP verbs
- Interpretation is subject to multiple processes instead of a single next-mention or production bias

Null pronoun

None of the model explains well

- Interpretation is always highly N1-biased
- Less sensitive to the verb bias, and none of the model can capture the sensitivity well

Thank you!

References

Ariel, M. (1990). *Accessing noun-phrase antecedents*: Routledge.

Arnold, J. E. (2001). The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. *Discourse processes, 31*(2), 137-162. doi:10.1207/s15326950dp3102_02

Baayen, R. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511801686

Givón, T. (1983). Topic continuity in discourse: An introduction. *Topic continuity in discourse: A quantitative cross-language study, 3*, 1-42.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 274-307. doi:doi:10.2307/416535

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and Coreference Revisited. *Journal of semantics, 25*(1), 1-44. doi:10.1093/jos/ffm018

Kehler, A., & Rohde, H. (2019). Prominence and coherence in a Bayesian theory of pronoun interpretation. *Journal of Pragmatics, 154*, 63-78. doi:10.1016/j.pragma.2018.04.006

Patterson, C., Hagen, J., Kehler, A., & Schumacher, P. B. (2020). *A Bayesian approach to modelling German personal and demonstrative pronouns*. Poster presented at the AMLaP 2020, University of Potsdam, Germany.

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience, 29*(8), 912-927. doi:10.1080/01690965.2013.854918

References

Ueno, M., & Kehler, A. (2016). Grammatical and pragmatic factors in the interpretation of Japanese null and overt pronouns. *Linguistics*, *54*(6), 1165-1221.

Zhan, M., Levy, R., & Kehler, A. (2020). Pronoun interpretation in Mandarin Chinese follows principles of Bayesian inference. *PLoS One, 15*(8), e0237012.

Zhang, A. (2018). *Interpretation biases and processing of null and overt pronouns in Chinese*. Department of English Language and Literature. Unpublished Doctoral Dissertation. Konkuk University.

Additive smoothing

- Pseudo-count one occurrence for each referential form (Names, overt, null) under each possible referent

Next-mention bias without smoothing

$$\frac{Count(\text{referent})}{Count(\text{Subject}) + Count(\text{object})}$$

Production bias without smoothing

$$\frac{Count(\text{referent ^ pronoun})}{Count(\text{referent})}$$

Next-mention bias with smoothing

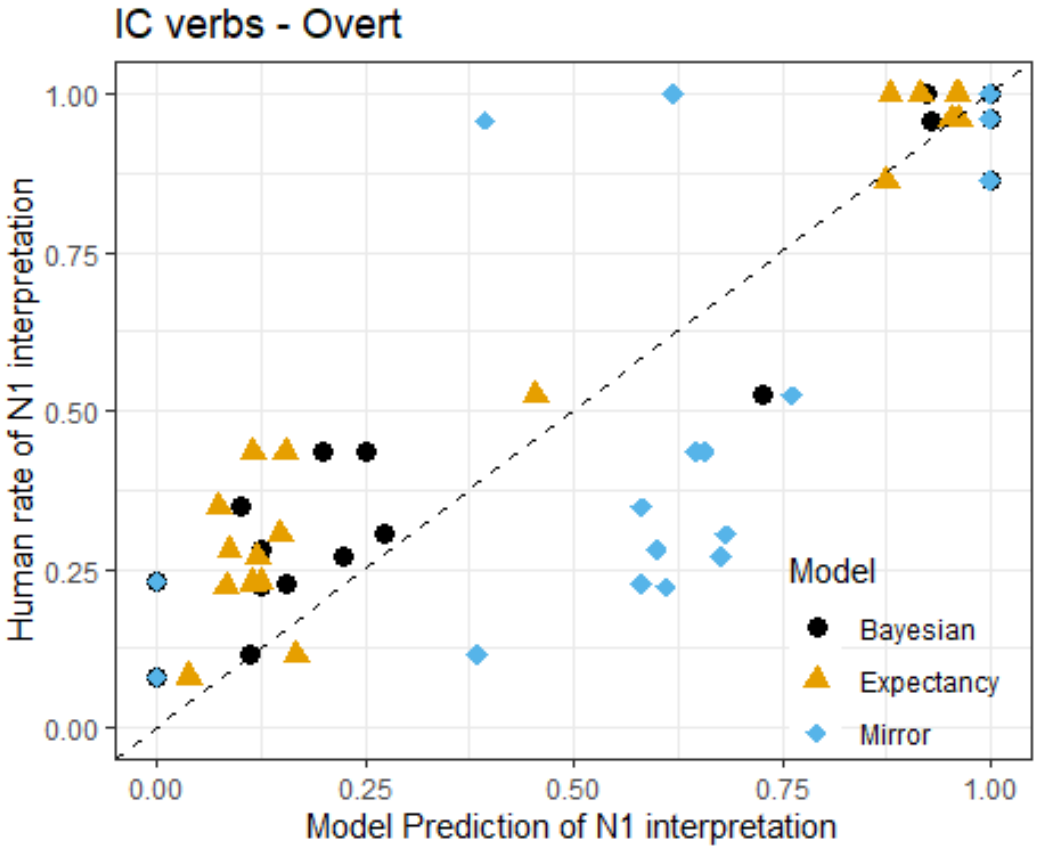
$$\frac{Count(\text{referent}) + 3}{Count(\text{Subject}) + Count(\text{object}) + 2 \times 3}$$

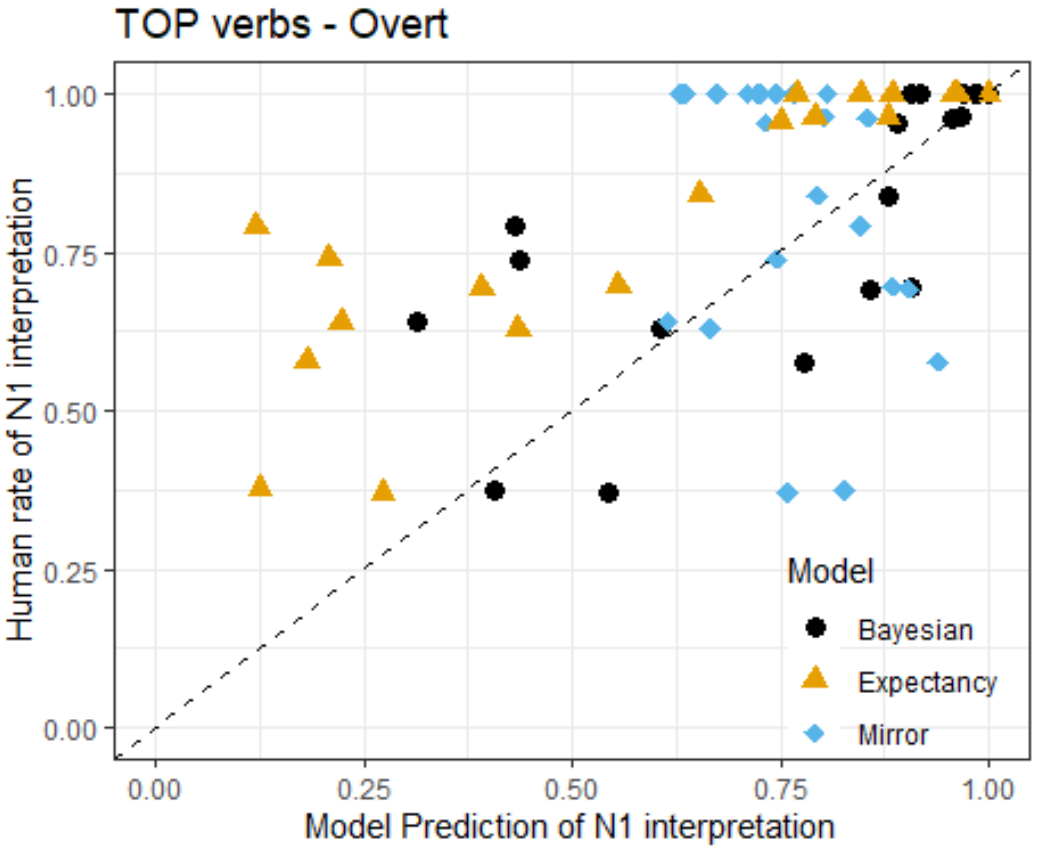
Production bias with smoothing

$$\frac{Count(\text{referent ^ pronoun}) + 1}{Count(\text{referent}) + 3}$$

Model Evaluation – Overt pronoun

Without smoothing





Model Evaluation – Null pronoun

Without smoothing