

Mentioning Atypical Properties of Objects is Communicatively Efficient

Elisa Kreiss¹, Judith Degen², Robert X. D. Hawkins², Noah D. Goodman²

¹University of Osnabrueck, ²Stanford University

Introduction

- **Problem of content selection:** which factors determine the information that speakers include in referring expressions?¹

1.(over)informativeness²

- say “*blue* banana” when only one banana

2.cost

- less likely to say long or infrequent property

3.typicality^{3,4,5}

- “*blue* banana” more likely than “*yellow* banana”

- Unified account of overinformative referring expressions lacking

- **Our approach:** when should a *rational* speaker mention an object’s color?


Typicality norming study

- Collect empirical typicality values for each utterance-object pair

- 3 separate studies

1. adjective + noun (“blue banana”?)
2. noun (“banana”?)
3. adjective (“blue”)

How typical is this object for a **banana**?




very atypical

very typical

Continue

How typical is this color for the color **yellow**?



very atypical

very typical

Continue

Results

Example typicality values for the banana case; numbers shown in bold are “correct” pairings.

Utterance	Banana items			Other
	yellow	brown	blue	
<i>banana</i>	.98	.66	.42	.05
<i>yellow banana</i>	.98	.33	.17	.05
<i>brown banana</i>	.28	.90	.18	.04
<i>blue banana</i>	.20	.18	.91	.06

Conclusion

- Speakers redundantly mention color when confusability of intention is otherwise high
- RSA with continuous semantics captures this
- ~~overinformative~~ referring expressions
 - ↑
- rationally redundant

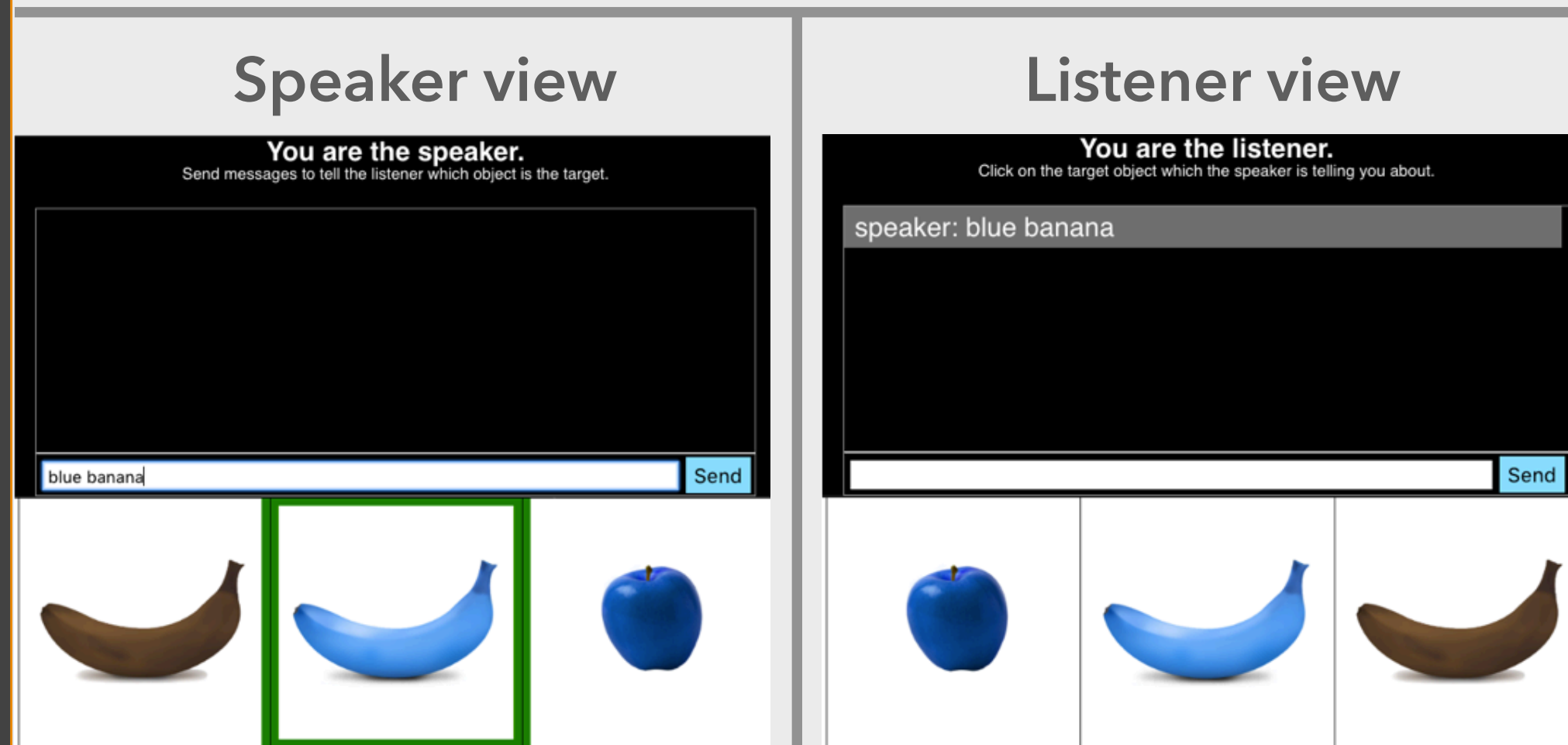
References

- ¹ Grice, H. P. (1975). Logic and Conversation. *Syntax and Semantics*, 3 , 41-58.
² Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, 27(1), 89-110.
³ Rubio-Fernandez, P. (2016). How redundant are redundant color adjectives? An efficiency-based analysis of color overspecification. *Frontiers in Psychology*, 7 (153).
⁴ Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: evidence for effects of informativity expectations. *Journal of psycholinguistic research*, 32(1), 3-23.
⁵ Westerbeek, H., Koolen, R., & Maes, A. (2015). Stored object knowledge and the production of referring expressions: the case of color typicality. *Frontiers in Psychology*, 6(July), 1- 12.
⁶ Goodman, N.D. & Frank, M.C. (2016). Pragmatic language interpretation as pragmatic inference. *Trends in Cognitive Sciences*. 20(11), 818-829.

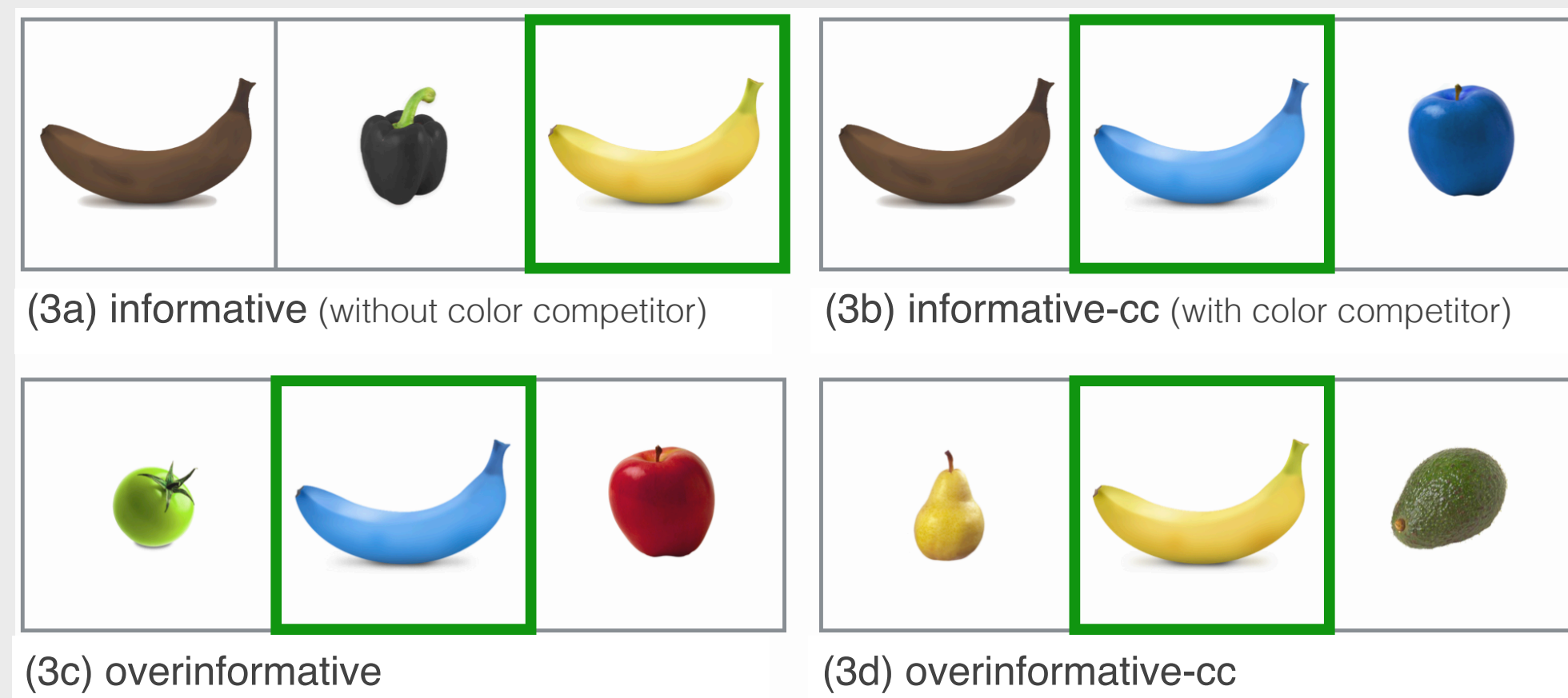
Production study

- Collect freely produced referring expressions through chatbox in two-player reference game
 - *Speaker* aim: Communicate target
 - *Listener* aim: Click on target

- Recruited 60 pairs on Amazon Mechanical Turk

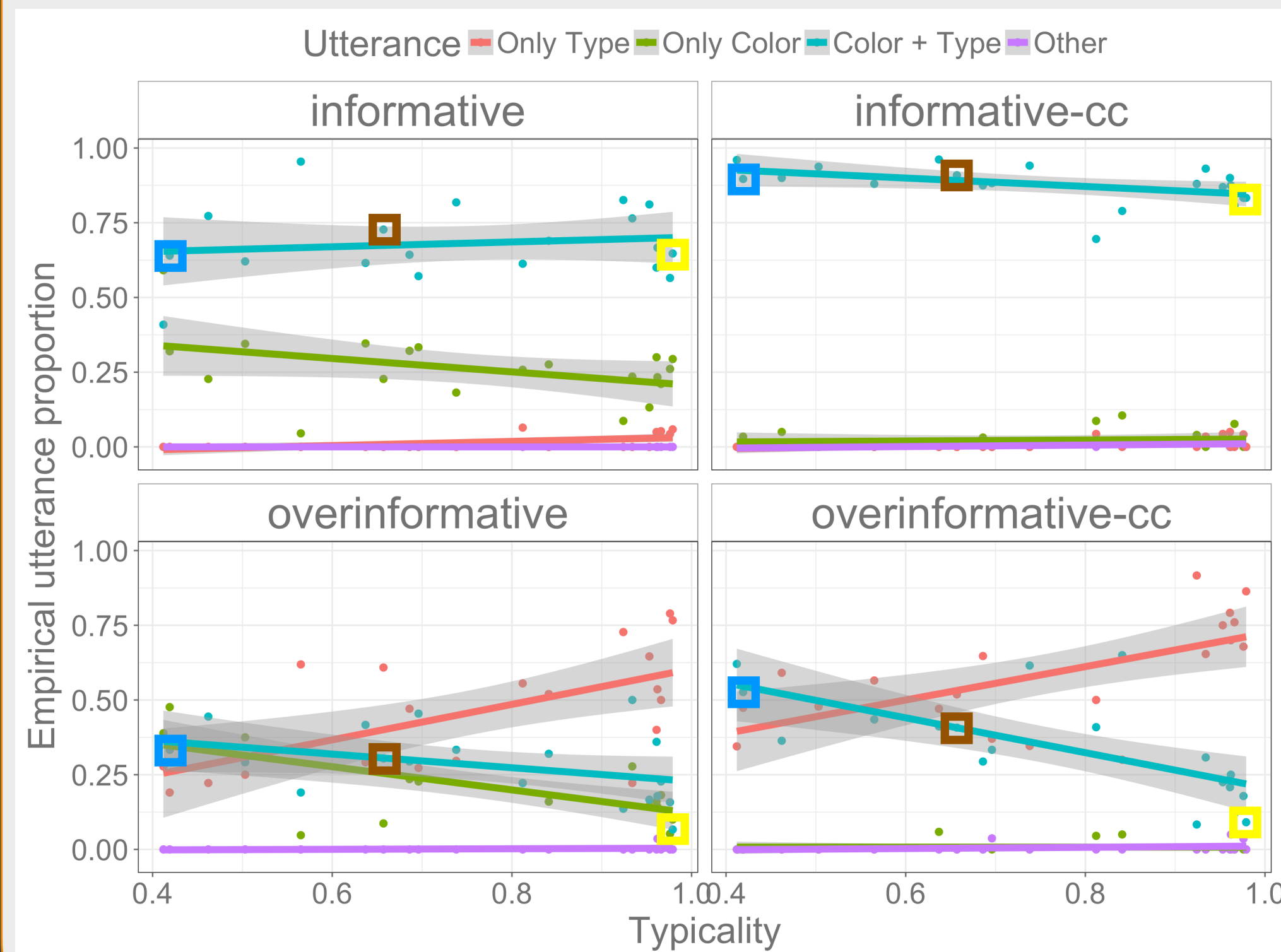


Conditions



Results

- Replicated typicality effect in overinformative conditions (more overinformative mentions of atypical colors)
- Also found typicality effect in informative conditions (“COLOR banana” cases marked)



Computational model

- Formalize in Rational Speech Act (RSA) framework⁶
- Literal listener L_0 selects between contextual referents lexicon according to lexicon \mathcal{L} :

$$L_0(c|u, C) \propto \exp(\lambda_{typ} \mathcal{L}(u, c))$$

- Pragmatic speaker S_1 selects utterance to communicate an intended referent c_i by trading off *informativity* with *cost*:

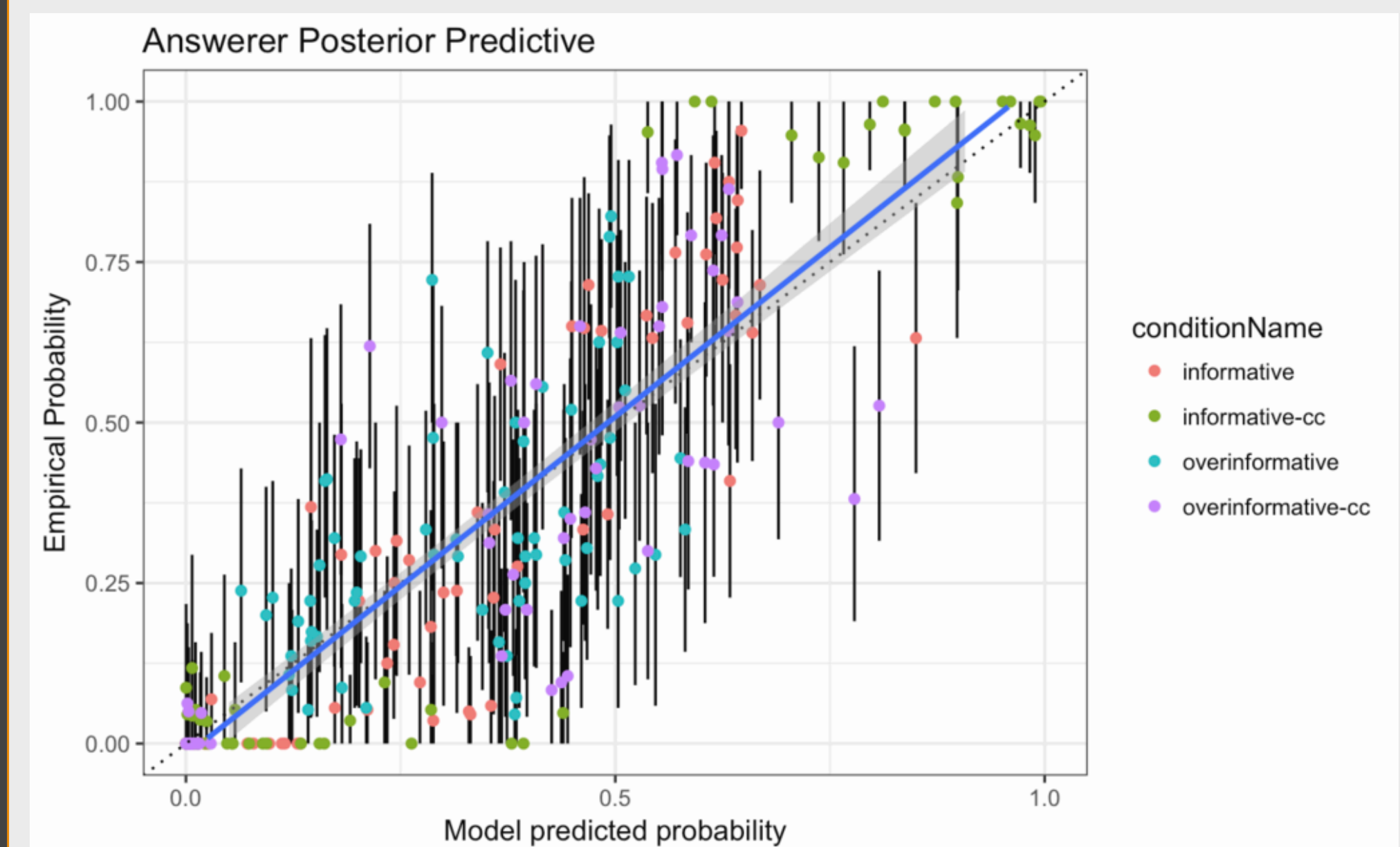
$$S_1(u|c_i) \propto \exp(\alpha \log(L_0(c_i|u, C)) - \text{cost}(u))$$

- Cost is defined as a function of an utterance’s length and its corpus frequency

$$\begin{aligned} \text{cost}(u) = & \beta_{freq} \hat{c}_f + (1 - \beta_{freq}) \hat{c}_l + \\ & \beta_{adj} \delta_{adj \in u} + \\ & \beta_{noun} \delta_{noun \in u} \end{aligned}$$

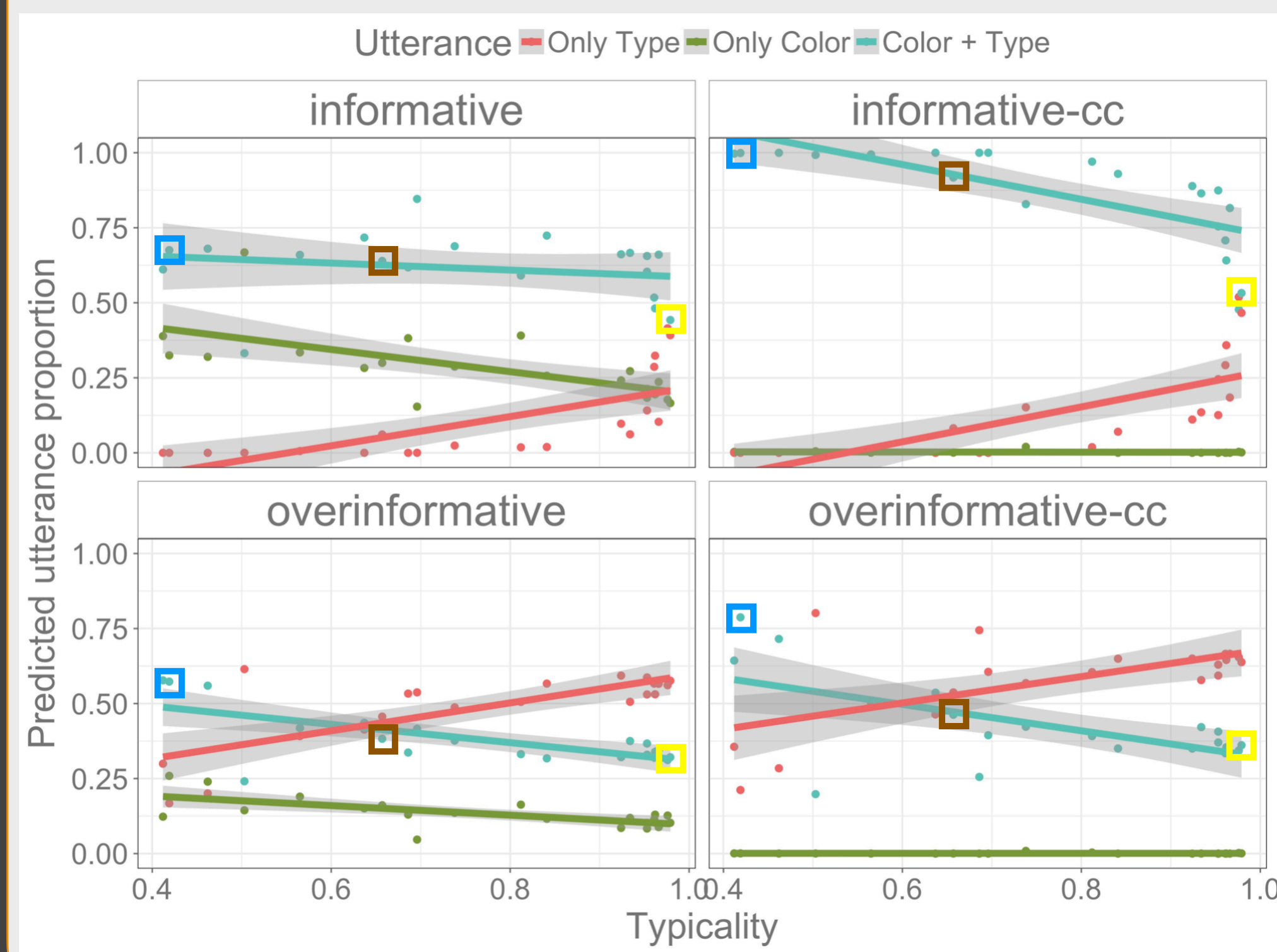
- Critically, we use a real-valued lexicon

$$\mathcal{L}(u, c) = \llbracket u \rrbracket(c)$$



Correlation between empirical and predicted utterance probability; $R^2 = .75$

(“COLOR banana” cases marked)



Discussion & Problems

informative and overinformative condition work against each other: in the informative condition the color and color-and-type expressions need to show a bigger gap; in the overinformative condition the gap needs to be closed

alternative: noise-model (either replacing or adding objects of the same type or color to context to achieve uncertainty about objects in context)

Acknowledgments



