Mentioning Atypical Properties of Objects is Communicatively Efficient

Elisa Kreiss[1], Judith Degen, Robert X. D. Hawkins, Noah D. Goodman[2]

University of Osnabrueck[1], Stanford University[2]

Introduction

What governs how much information speakers include in referring expressions?

One important pressure is for speakers to include just enough information for their interlocutor to uniquely select an intended referent from among a set of potential referents (say “yellow banana” in Fig 3a and only “banana” in Fig 3c). However, speakers have a well-documented preference to mention properties “overinformatively”, especially color. E.g., speakers are likely to call the target in 3c a “blue banana”. More precisely, speakers tend to mention atypical rather than typical properties of objects overinformatively. For example, a banana is more likely to be called a “blue banana” if it is blue and more likely just to be called a “banana” if it is yellow.

A unified account of this phenomenon is still lacking. We ask: when *should* a rational speaker mention an object’s color?

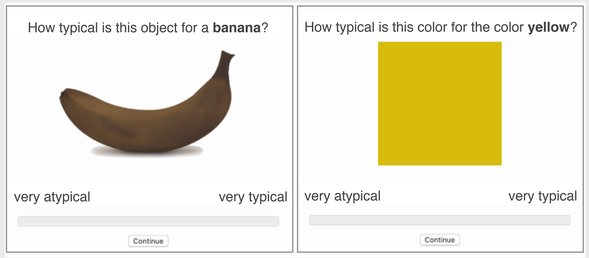
Typicality norming studies

Purpose

- Collect empirical typicality values for each utterance-object pair

Design

- 3 separate studies
  1. adjective&noun-object pairs
  2. noun-object pairs
  3. adjective-color pairs

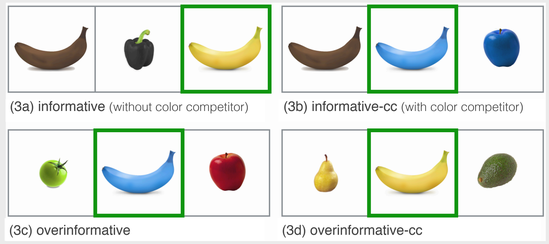


Examples for noun-object and adjective-color trials.

Result

Utterance	Banana items: yellow	brown	blue	Other
banana	**.98**	**.66**	**.42**	.05
yellow banana	**.98**	.33	.17	.05
brown banana	.28	**.90**	.18	.04
blue banana	.20	.18	**.91**	.06

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



The four context conditions as they occurred in the study.

Referential expression study

Purpose

- Collect freely produced referring expressions

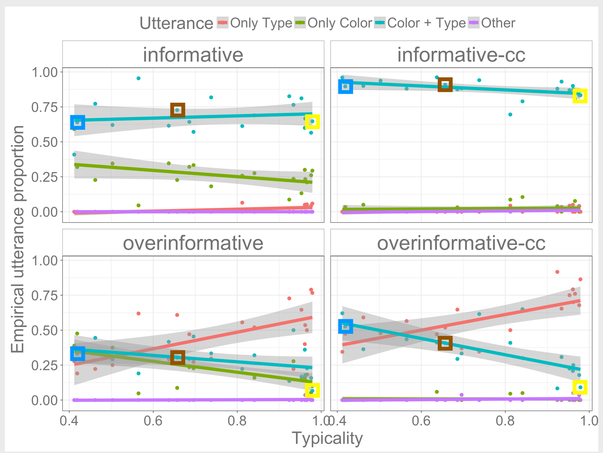
Design

- Two-player reference game on MTurk
- 60 pairs
- Speaker aim: Communicate target
- Listener aim: Correctly identify it
- Free communication through chat box

Experimental design.

Results

- Replicated typicality effect in overinformative conditions
- Also found typicality effect in informative conditions



Empirical utterance proportions with respectively marked “COLOR banana” cases.

Model

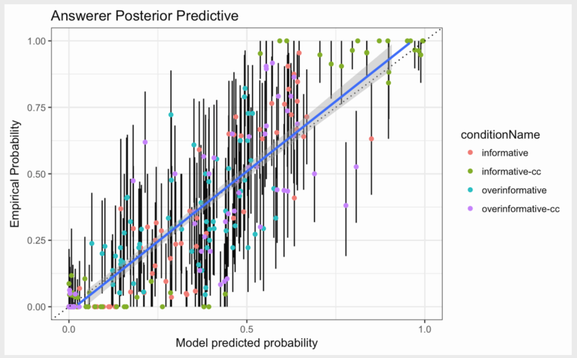
- Rational Speech Act (RSA) framework
- Literal listener L_0 that selects between contextual referents c ∈ C proportionally to the meanings given by a lexicon L:

$$L_0(c|u,\mathcal{C}) \propto \mathcal{L}(u,c)P(c)$$

- Assume uniform prior beliefs P(c) over referents
- Pragmatic speaker S_1, which selects an utterance $u \in U$ to communicate an intended referent c_i by trading off *informativity* with *cost*:

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

- Cost is usually defined as a function of an utterance’s length or corpus frequency
- α is a parameter controlling the speaker’s “optimality”



typicality weights and allowing costs to go negative; R^2 = .75, I think, no noise; soft semantics

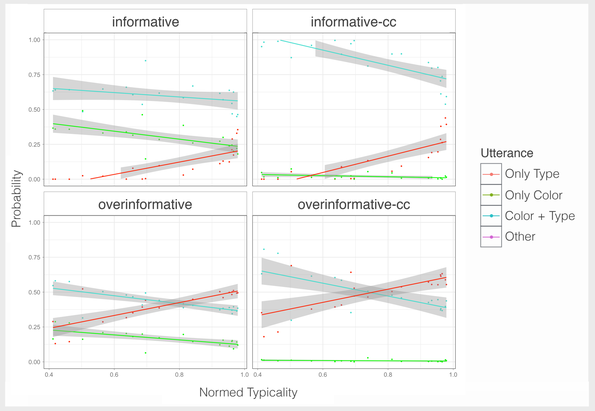


Figure 7. bla

Conclusion

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

References

Acknowledgments