On the purpose of ambiguous utterances

Gregory Scontras[1], Asya Achimova[2], Christian Stegemann[2], Martin Butz[2]

[1]University of California, Irvine; [2]University of Tübingen

a feature, not a bug

Traditionally, linguists have treated **ambiguity** as **a bug in the communication system**, something to be avoided or explained away (Grice, 1975; Chomsky, 2002).

More recent research has begun to take notice of the **efficiency** ambiguity affords to us: by relying on context to fill in missing information, we can **reuse lightweight bits of language** rather than fully specifying the intended message (Levinson, 2000; Piantadosi et al., 2012; Wasow, 2015).

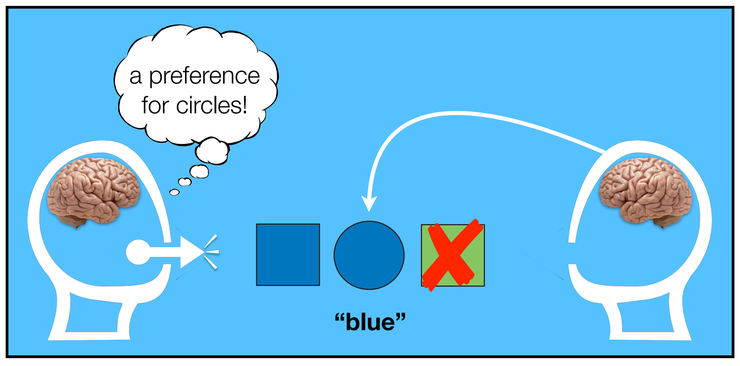
Viewed in this way, ambiguity serves as a feature, not a bug, of an **efficient communication system**.

The current work identifies **an additional benefit** in using ambiguous language: the **extra information** we gain from observing how our listeners resolve ambiguity.

We propose that **language users learn about each other's private knowledge by observing how they resolve ambiguity**.

simple reference games

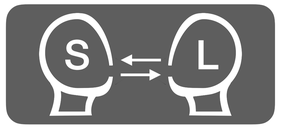
If language does not do the job of specifying the information necessary for full interpretation, then listeners are left to draw on their opinions, beliefs, and preferences to fill in the gaps; **by observing how listeners fill those gaps, speakers learn about the opinions, beliefs, and preferences (i.e., the priors) of listeners**.



a computational model

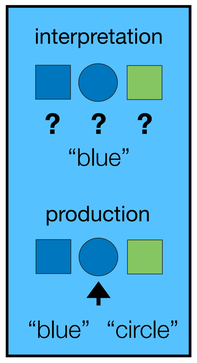
Rational **S**peech **A**ct

speaker observes state, chooses utterance
listener hears utterance, infers state



speaker and listener coordinate:
utterance + interpretation that maximizes the probability of correctly resolving the **Q**uestion-**U**nder-**D**iscussion

(Frank & Goodman, 2012; Goodman & Frank, 2016)



introducing preferences

We consider **two different models**.

One model is **fully pragmatic**, with listeners interpreting utterances by **reasoning about speakers:**

a **fully pragmatic** listener:

$$P_{L_1}(s \mid u, \boldsymbol{f}) \propto P_{S_1}(u \mid s) \cdot \boldsymbol{P(s \mid f)}$$

$$P_{S_1}(u \mid s) \propto exp(\alpha \cdot [\log(P_{L_0}(s \mid u)) - C(u)])$$

$$P_{L_0}(s \mid u) \propto [\![u]\!](s)$$

The other model is **simpler**, with listeners interpreting utterances **without invoking** reasoning about **the speaker:**

a **simpler** listener:

$$P_{L_0}(s \mid u, \boldsymbol{f}) \propto [\![u]\!](s) \cdot \boldsymbol{P(s \mid f)}$$

Recent work (Sikos et al., 2019) suggests that **simpler models might be better** suited to handle reasoning **in simple reference games**.

inferring preferences

The speaker **observes** the listener's **object choice**, then **infers the preferences** that led to the choice.

$$P_{S_2}(\boldsymbol{f} \mid s, u) \propto P_{L_1}(s \mid u, f) \cdot P(f)$$

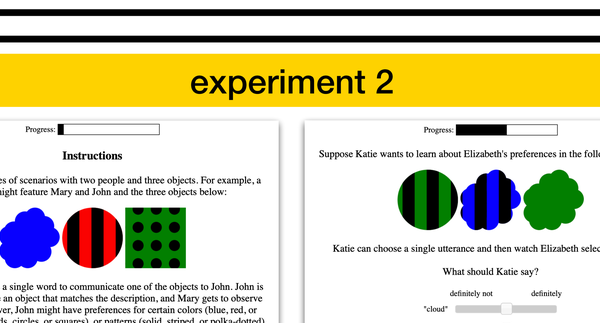
$$P_{L_1}(s \mid u, \boldsymbol{f}) \propto P_{S_1}(u \mid s) \cdot \boldsymbol{P(s \mid f)}$$

$$P_{L_0}(s \mid u, \boldsymbol{f}) \propto [\![u]\!](s) \cdot \boldsymbol{P(s \mid f)}$$

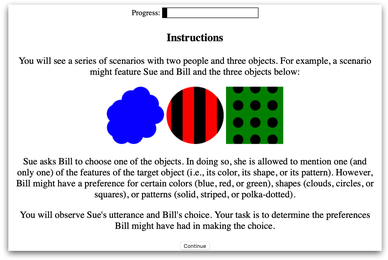
choosing utterances

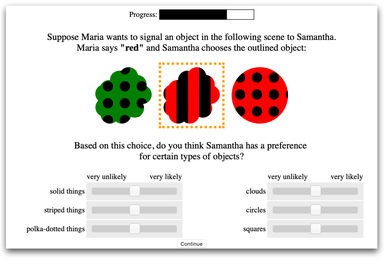
Useful utterances **maximize information gain**; they maximize the **difference between** the **prior** and the **posterior**.

$$P_b(u) \propto \sum_{s:\,[\![u]\!](s)} \boldsymbol{\lambda} \cdot \mathrm{KL}(P(f), P_{S_2}(f \mid s, u)) - C(u)$$

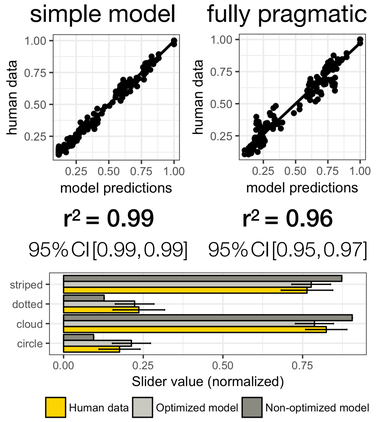


experiment 1



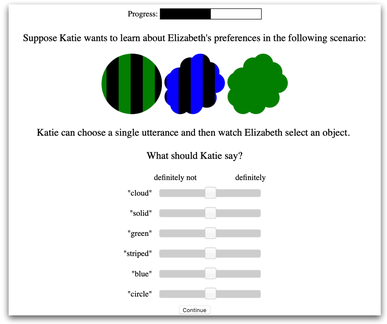


- **82 native speakers** of English
- Followed Frank & Goodman 2012 in stimuli creation
- **15 trials**: 10 potentially informative about preferences, 5 uninformative
- **Preference strength** ($P(s \mid f)$) **fit to** individual participant **data**
  - **15** participants indicated **weak preferences**, suggesting lack of engagement with task
  - **67** participants indicated **strong preferences**
- **Model performs well** even **with default parameters**

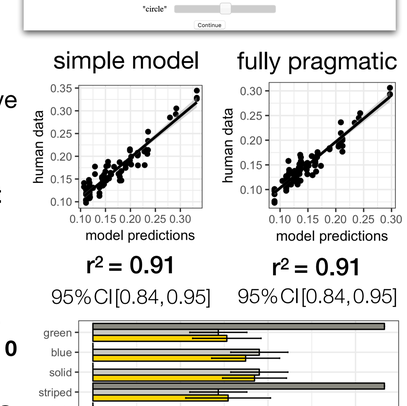


non-optimized r^2: 0.97; 95% CI [0.96, 0.97]

experiment 2



- **82 native speakers** of English
- Object scenes from Experiment 1
- **15 trials**: 10 potentially informative about preferences, 5 uninformative
- **Sensitivity to informativity** (λ) **fit to** individual participant **data**
  - **9** participants had λ **close to 0**
  - **33** participants had λ **less than 0**
    - Prefer **unambiguous** utterances
  - **40** participants had λ **greater than 0**
    - Prefer **ambiguous** utterances
- **Three** different participant **groups**



non-optimized r^2: 0.07; 95% CI [0.01, 0.23]

Optimization has a **large effect!**

References: **Chomsky (2002)**. An interview on minimalism. In A. Belletti & L. Rizzi (Eds.), *On nature and language* (p. 92-161). Cambridge: Cambridge University Press. **Frank & Goodman (2012)**. Predicting pragmatic reasoning in language games. *Science*, *336*, 998- 998. **Goodman & Frank (2016)**. Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, *20*(11), 818-829. **Grice (1975)**. Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (p. 26-40). New York: Academic Press. **Levinson (2000)**. Presumptive meanings: The theory of generalized conversational implicature. Cambridge, MA: MIT Press. **Piantadosi et al. (2012)**. The communicative function of ambiguity in language. *Cognition*, *122*, 280-291. **Sikos et al. (2019)**. Reevaluating pragmatic reasoning in web-based language games. Poster presented at The 8th Experimental Pragmatics conference, University of Edinburgh. **Wasow (2015)**. Ambiguity avoidance is overrated. In S. Winkler (Ed.), *Ambiguity: Language and communication* (p. 29-47). de Gruyter.