# The added informativity of ambiguous utterances

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

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

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. This reasoning accords with years of psycholinguistic research documenting that speakers readily produce ambiguous utterances (e.g., Ferreira, 2008). 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. 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 of the listeners.

Our model of ambiguous language builds on the vanilla Rational Speech Act model of Frank and Goodman (2012) by allowing for uncertainty around the listener's state prior, $P(s)$. We have in mind a scenario where a listener might have a preference for a certain object feature (e.g., blue things, squares, circles, etc.), and these preference will influence their object choice. With this in mind, the speaker produces an utterance $u$, observes the listener's referent choice $s$, and, on the basis of that choice, infers the preferences $f$ the listener might have had when making the choice. We use the same $L_0$ and $S_1$ from the vanilla model. However, we now parameterize $L_1$'s state prior so that it operates with respect to a given feature preference $P(s|f)$:

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

We then model a pragmatic speaker $S_2$ who updates beliefs about $L_1$'s preferences, $P(f)$. To do so, $S_2$ produces $u$, then observes $L_1$'s choice of $s$, and finally reasons about the likely feature preference $f$ that $L_1$ used to make that choice:

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

Moreover, we model the reasoning process by which a speaker selects the best utterance to learn about the preferences of the listener. Starting with no knowledge of the listener's preferences, $S_2$ can be assumed to expect a uniform (i.e., flat) feature preference prior $P(f)$. The more the speaker's posterior beliefs about the preferences, $P_{S_2}(f|u, s)$, deviate from the uniform prior, the more the speaker will have learned about the listener's preferences. We can thus model this reasoning in the light of expected information gain, which can be equated with the attempt to maximize the KL divergence between the speaker's flat prior and the expected posterior of the listener's feature preferences $f$, integrating over all hypothetically possible state observations $s$:

$$P_b(u) \propto \sum_{s:\ \llbracket u \rrbracket(s)} \lambda \cdot \mathrm{KL}(P(f), P_{S_2}(f|u, s)) - C(u).$$

**Expt. 1: Testing $S_2$ predictions.** We presented 82 English-speaking participants recruited via MTurk with a series of reference game scenarios modeled after Frank and Goodman

(2012). Each scenario featured two people and three objects. One of the people served as the speaker, and the other served as the listener. The speaker asks the listener to choose one of the objects, but in doing so she is allowed to mention only one of the features of the target object. Participants were told that the listener might have a preference for certain object features, and participants were tasked with inferring those preferences after observing the speaker's utterance and listener's object choice. Participants completed a series of 15 trials. Objects and utterances were chosen at random with the constraint that 10 trials were potentially informative with respect to listener preferences and 5 trials were uninformative (e.g., observing that the listener chose one of three identical objects).

To generate model predictions, we fixed the strength of individual feature preferences $f_i$ with a softness parameter $\gamma$, which determines how strong a preference is if a certain feature value is indeed preferred. We optimized $\gamma$ in the light of the KL divergence between the individual participants' slider values and the corresponding model predictions:

$$\text{KL} = \sum_{i=1}^{n} P(f_i'|u, s)(\log(P(f_i'|u, s) - \log(P(f_i|u, s)),$$

where $P(f_i'|u, s)$ specifies a participant's normalized slider value settings for a given stimulus and $P(f_i|u, s)$ specifies the respective model prediction. We observed a strong positive correlation between the human judgments and model predictions ($r^2 = 0.96$, 95% CI[.94,.97]). Thus, we find strong empirical support for our extended RSA model of preference inference: speakers are able to use listener behavior to arrive at information about their preferences.

**Expt. 2: Testing $P_b(u)$ predictions.** Participants (n=82) encountered a reference game scenario similar to Expt. 1 in which a speaker signals an object to a listener who might have a preference for certain types of objects. Rather than observing the utterance and referent choice, participants were now tasked with helping the speaker choose an utterance that was "most likely to reveal the listener's color, shape, or pattern preferences." Each trial featured a set of three objects. Participants adjusted sliders to indicate which single-feature utterance the speaker should choose. Potential utterances corresponded to the features present. To generate model predictions, we must also set the $\lambda$ parameter, which factors the importance of choosing the expected most informative utterance with respect to the determined KL divergence values. We used the same optimization procedure from Expt. 1, and we again found a strong positive correlation between human judgments and model predictions ($r^2 = 0.91$, 95% CI[.84,.95]). Thus, we find evidence in support of the idea that speakers reason pragmatically about the relative informativity of ambiguous language.

The results of Experiment 1 demonstrate that naïve speakers are able to reason pragmatically about *why* listeners may take the actions they do, and the success of our computational model in predicting the observed behavior offers an articulated hypothesis about *how* this reasoning proceeds: when speakers are aware of the ambiguity in their utterances, observing how listeners resolve that ambiguity provides clues to the preferences listeners use when doing so. The results of Experiment 2 demonstrate that speakers are able to capitalize on this reasoning to strategically select utterances that are most likely to inform their understanding of the preferences of their listeners, and that the most informative utterances are also the most ambiguous ones.

## References

Chomsky, N. (2002). An interview on minimalism. In A. Belletti and L. Rizzi (Eds.), *On Nature and Language*, pp. 92–161. Cambridge: Cambridge University Press.

Ferreira, V. S. (2008). Ambiguity, accessibility, and a division of labor for communicative success. *Psychology of Learning and Motivation: Advances in Research and Theory 49*, 209–246.

Frank, M. C. and N. D. Goodman (2012). Predicting pragmatic reasoning in language games. *Science 336*, 998–998.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech Acts*, pp. 26–40. New York: Academic Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.

Piantadosi, S. T., H. Tily, and E. Gibson (2012). The communicative function of ambiguity in language. *Cognition 122*, 280–291.

Wasow, T. (2015). Ambiguity avoidance is overrated. In S. Winkler (Ed.), *Ambiguity: Laguage and Communication*, pp. 29–47. de Gruyter.