Roses and flowers: an informativeness implicature in probabilistic pragmatics

Roger Levy (rlevy@ucsd.edu), Leon Bergen (bergen@mit.edu), Noah Goodman (ngoodman@stanford.edu)

In the half-century since the introduction of Grice's maxims (1957; 1975), considerable effort has gone into refining them into a smaller set of generalizations rooted in deeper principles of cooperative communication (Horn, 1984; Sperber and Wilson, 1986; Levinson, 2000, inter alia). One particularly fruitful result has been identification of the tension between "quantity" (Q-)implicature, in which utterance meanings are upper-bounded by the literal content of alternatives, as in (1), and "informativeness" (I-)implicature, in which utterances are interpreted as strengthened to a prototypical case, as in (2) (Atlas and Levinson, 1981; Horn, 1984; Levinson, 2000):

(1) a. Pat has three children→Pat has exactly three children
b. I injured a child yesterday→The child was not mine

(2) I injured a finger yesterday→The finger was mine

A Bayesian account to pragmatic inference offers the promise of this tension falling out of more general principles: complex interactions are predicted from recursive reasoning involving alternative utterances, shared beliefs about common communicative goals, prior information about world state, and utterance costs. Here we discuss the challenges posed to such an account by a previously unobserved pattern of informativeness implicature: when the conjunction of a superordinate category X with a subordinate member x of that category, *x and X*, receives a strengthened interpretation equivalent to *x and other X*, as in (3) below:

(3) We sell roses and flowers for Mother's Day.[1]

Corpus analysis shows that English has many such common alternations: *tulips and (other) flowers*, *beef and (other) meat*, *horse and (other) animal*, *physicists and (other) scientists*, and more. Longitudinal data show that this is an historically stable pattern. An experimental investigation of naive native speaker intuitions about how many flower types are being talked about shows that omitting *other* has no discernible effect on interpretation.

The challenge for a formal analysis is thus to show how *roses and flowers* can come to be interpreted as meaning the same thing as *roses and other flowers*. The immediate challenge for a strongly neo-Gricean account is that if literal semantics have a chance to be computed globally, we are stuck with a truth-conditional meaning for utterances involving *roses and flowers* that is the same as for utterances involving *roses* alone: for example, the literal meaning of (4) is the same as that of (3):

(4) We sell roses for Mother's Day.

The strengthening of (3) would need to be a "division of pragmatic labor", with the more formally marked of a pair of literally meaning-equivalent expressions associated with more unusual meanings than the less marked (Horn, 1984; Levinson's M-implicature). Problematically, however, this would predict that *roses and flowers* could be strengthened to mean *roses and no other flowers* in cases when such a state of affairs is more unusual (has lower prior probability) than *roses and other flowers*. As a second challenge, if *roses and other flowers* is considered an alternative utterance (seemingly necessary to yield the more familiar Q-implicature of *John bought roses* that John bought no other types of flowers), it is not clear why *roses and flowers* fails to trigger a Q-implicature of *roses and no other flowers*.

[1] `http://e-clubhouse.org/sites/townofsheboyganwi/`

Here we present a rational speech-act theory (Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013; see also Jäger, 2012; Franke, 2013) account of this pragmatic strengthening that is robust to precise details of prior probabilities and specification of alternative utterances. We model the listener-speaker relationship as a pair of recursive probabilistic functions, with listeners as rational Bayesian interpreters and speakers as soft-max rational actors. The set of possible world states is given in Figure 1, with the literal semantic content of each simple NP expression outlined (f_1 being roses).

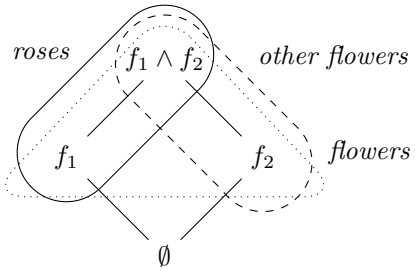

Figure 1: The domain of possible flower types

We employ the technique of LEXICAL UNCERTAINTY first introduced by Bergen et al. (2012) to account for M-implicature by introducing explicit reasoning over different possible mappings between forms and pragmatically refined meanings, allowing the efficient pairing of low-cost forms with high-probability meanings to be identified and subsequently strengthened through recursive inference. Speakers' and interpreters' reasoning follows these recursive probabilistic functions:

$$P^{(0)}_{Listener}(m|u,\mathcal{L}) \propto \mathcal{L}_u(m)P(m)$$

$$P^{(1)}_{Speaker}(u|m,\mathcal{L}) \propto \left[P^{(0)}_{Listener}(m|u,\mathcal{L})e^{-c(u)}\right]^{\lambda} \quad P^{(1)}_{Listener}(m|u) \propto P(m)\sum_{\mathcal{L}} P(\mathcal{L})P^{(1)}_{Speaker}(u|m,\mathcal{L})$$

$$P^{(n)}_{Speaker}(m|u) \propto \left[P^{(n-1)}_{Listener}(m|u)e^{-c(u)}\right]^{\lambda} \qquad P^{(n)}_{Listener}(m|u) \propto P(m)P^{(n)}_{Speaker}(u|m) \qquad (n>1)$$

for meanings m, cost function c, utterances u, greedy optimality parameter λ, and ranging over lexica $\mathcal{L}$—refinements of the "literal" form-meaning mappings of Figure 1. Accounting for *roses and flowers* requires lexical uncertainty to be COMPOSITIONAL: the form-meaning mappings for simple NP expressions (*roses*, *flowers*, *other flowers*) can be refined arbitrarily but complex expressions (*roses and flowers*, *roses and other flowers*) must mean the composition of the refined meanings of their constituent parts.

This model robustly recovers the empirically observed strengthening for *roses and flowers*. In addition, it makes a distinctive prediction regarding how speakers' preferences regarding *other*-drop should vary as a function of the prototypicality of the distinguished subtype f_1—modeled here as $P(f_1|\textit{flowers})$ for Figure 1. As $P(f_1|\textit{flowers})$ increases, an increasingly strong M-implicature bias is added to the compositional model's fundamental bias toward the empirically observed strengthening. The model thus predicts that *other*-drop will be more frequent the more prototypical f_1 is in the supertype. Using unigram word frequency as a proxy for in-category prototypicality, we find support for this prediction in corpus counts (Google Web n-grams) of expressions of the type *x and flowers*, *flowers and x*, and *x and other flowers*. As seen in Figure 2, across a variety of flower types x, higher unigram frequency of x is associated with higher rates of *other* drop.

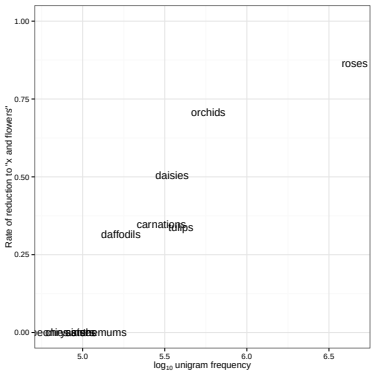

Figure 2: Rate of *other*-reduction as a function of corpus frequency

In sum, this work is the first report of a new class of informativity implicature and shows how simple principles of rational communication can explain its major patterns of both interpretation and speaker choice.