July 2015. To appear in *Proceedings of BLS 41*.

Negotiating Lexical Uncertainty and Speaker Expertise with Disjunction

CHRISTOPHER POTTS AND ROGER LEVY
*Stanford University and UC San Diego**

Abstract

There is a well-known preference for disjunctions X *or* Y to be construed so that X and Y are semantically disjoint. However, there are two felicitous usage patterns in which the speaker violates this preference in part to convey information about the language itself. First, disjunctions of terms in a one-way semantic inclusion relation, such as *boat or canoe*, can form part of a speaker strategy to manage lexical uncertainty surrounding the two terms, or block unwanted implicatures that the listener might draw from the general term alone. Second, disjunctions of synonymous terms like *wine lover or* ***oenophile*** can be used to convey definitional information. We explore both of these uses, relying on corpora to obtain a fuller picture of their motivations and their effects on the listener. In addition, we show how both these uses are predicted by a standard semantics for disjunction and a recursive probabilistic model of communication in which speakers and listeners simultaneously exchange information about the world and about the language they are using. We also use the model to begin to formally characterize the pragmatics of implicature cancelation or blocking.

1 Communicating in Language about Language

Natural languages are neither fixed across time nor identically reproduced in all speakers, but rather continually renegotiated during interactions (Clark 1997). Discourse participants accommodate to each other's usage patterns (Giles et al. 1991), form temporary lexical pacts to facilitate communication (Clark & Wilkes-Gibbs 1986; Brennan & Clark 1996), and instruct each other about their linguistic views. Some of this communication in language about language is direct, as with explicit definitions like *'oenophile' means 'wine lover'*, but much of it arrives via secondary pragmatic inferences, as when X *such as* Y conveys that X subsumes Y (Hearst 1992; Snow et al. 2005).

Disjunction supports what appear to be opposing inferences about language. On the one hand, X *or* Y tends to convey that the meanings of X and Y are presumed to be disjoint (Hurford 1974), because the speaker holds such a view of the lexicon or is worried that the listener might. This pressure to EXCLUSIVIZE is robust enough to overcome even seemingly non-negotiable aspects of the lexicon; a medical webpage warns "If you still have symptoms or severe blockage in your arteries, you may need angioplasty or surgery", sending a clear signal that angioplasty and surgery are distinct options. Its continuation presupposes just that: "Having one of these procedures may save your leg". The disjunction might seem to

*For help with the data and ideas, we thank Arto Anttila, James Collins, Ashwini Deo, Itamar Francez, Kumiko Fukumora, Larry Horn, Justine Kao, Dan Lassiter, Natalia Roudakova, Rob Voigt, and Eva Wittenberg, as well as audiences at the 2015 LSA Annual Meeting and ESSLLI 2014. This research was supported in part by NSF IIS 1159679 and ONR N00014-13-1-0287 (CP), and an Alfred P. Sloan Foundation Research Fellowship and a Fellowship at the Center for Advanced Study in the Behavioral Sciences (RL).

be a needlessly verbose way of conveying the meaning of the more general disjunct, but the costs could be worth paying in virtue of the lexical side-effect of exclusivization.

In apparent opposition to exclusivization, disjunctions like *wine lover or oenophile* can be used to convey that the two disjuncts are roughly synonymous (Horn 1989), thereby providing secondary information that maximally violates the pressure to exclusivize. This inference is more elusive than the exclusivization inference, but it can arise in a broad range of contexts in which such DEFINITIONAL or IDENTIFICATIONAL information has social or communicative value, and there are stable orthographic and prosodic devices that help signal it. It is striking that both the definitional and exclusivization inferences, which seem so opposed, are supported by a single lexical item, and the puzzle deepens when we see that the empirical picture is not a quirk of English, but rather one found in a wide range of typologically and geographically diverse languages.

In this paper, we capture both of these classes of inference within a single recursive Bayesian model of pragmatic reasoning. The model finds its conceptual origins in Lewis's (1969) work on signaling systems and builds on ideas from iterated best response models (Jäger 2007, 2012; Franke 2009) and more thoroughly probabilistic variants of them (Camerer et al. 2004; Frank & Goodman 2012; Russell 2012). The crucial feature of our model is that it lets discourse participants communicate, not just about the world, but also about the language they are using (Bergen et al. 2012, 2014). From the speaker's perspective, this means that one's intentions in production are characterized in terms of both world information and linguistic information. From the listener's perspective, this means that pragmatic inference is cast as a problem of joint inference about the speaker's intended meaning and the speaker's preferred lexicon (Smith et al. 2013). We show that, within this model, both exclusivization and definitional inferences arise naturally from the expected semantic content of disjunction, depending on contextual parameters relating to speaker expertise, listener malleability, and information in the common ground. The model thus offers a genuinely pragmatic account of these inferences as well as characterizations of their stability and communicative value.[1]

2 Lexical Side-Effects from Disjunction

This section explores the exclusivization and definitional uses of disjunction. Our goal is to more precisely characterize what the inferences are like and to begin to understand which contexts steer speakers and listeners toward one or the other. These findings inform the modeling we describe in Sections 3–5.

2.1 Hurfordian Perceptions and Intentions

Hurford's (1974) generalization (HG) is a direct statement of the overall communicative pressure to treat disjuncts as exclusive:

(1) "The joining of two sentences by *or* is unacceptable if one sentence entails the other; otherwise the use of *or* is acceptable." (p. 410)

[1]Implementations of our model and related models, and all the code and data used in this paper, are available at `https://github.com/cgpotts/pypragmods/`.

The generalization is stated in terms of sentences, but Hurford's examples, given in (2) with his original judgments, make it clear that he intends it to hold for sub-sentential disjuncts as well:

(2) a. Ivan is an American or Russian.
b. The painting is of a man or a woman.
c. The value of x is greater than or equal to 6.
d. * John is an American or Californian.
e. * The painting is of a man or a bachelor.
f. * The value of x is greater than or not equal to 6.

Hurford uses HG to probe the nature and distribution of conversational implicatures (see also Gazdar 1979; Chierchia et al. 2012). Singh (2008) extends it to certain cases in which the disjuncts are merely overlapping. We endorse the guiding insight behind these accounts but reject the assumption that HG violations reliably lead to, or even correlate with, unacceptability or ungrammaticality. Disjunctions of apparently entailing phrases are routine (Simons 2001); all of the disjunctions marked as ungrammatical in (2) are found in fluent English text on the Web:

(3) a. "... and we trust that some of our American or Californian friends will tell us something of its growth of flower and fruit in its native habitats"
b. "It doesn't matter if you ask a boy or a man or a bachelor or even a husband"
c. "... the effect was greater than, or not equal to, the cause."

Russell (2012: §5.3) reports a numbers of similar Web-derived examples. Here is a sample:

(4) a. "We also rent only the most modern limos to our customers, because we believe that when you look for a limo service in Northern California or San Francisco, you want the best limousine service possible."
b. "By the time I've gone in I've had to pull out an animal or a cat that's on the verge of dying."
c. "Every now and again, people tend to change their surroundings. We update wall colors, change the drapes. Have new flooring installed. Sometimes we purchase new furniture or chairs."

We have collected a large corpus of apparent HG violations, available at the website for this paper. Here is a small sample from that corpus:

(5) a. "Stop discrimination of an applicant or person due to their tattoos."
b. "Promptly report any accident or occurrence."
c. "The anchor will lie on the bottom and the canoe or boat will be held by the stream's current."
d. "As an actor or performer, you are always worried about what the next job's going to be," Hensley says.
e. "After the loss of the animal or pet, there are further coping strategies available for the grieving individual."

f. "Bush was captured slyly removing a candy or gum from his mouth."

g. "Heroic is not a word one uses often without embarrassment to describe a writer or playwright . . . "

h. "But he never attended school during his senior year, never attended a party or prom."

The dataset includes 90 cases where the left disjunct entails the right, and 79 in which the right entails the left. However, we caution against using these counts to make inferences about general frequency or the relative prevalence of the two disjunct orders. We created the corpus using heuristic techniques based on WordNet (Fellbaum 1998) and ad hoc Web searches, so it can provide only a glimpse of what is possible. In addition, we have found that, for any two nouns N_1 and N_2 one believes to be in an overlap or proper entailment relation, it is generally possible to find contexts in which "N_1 or N_2" and "N_2 or N_1" are felicitous, and Web searches will generally yield examples.

Of course, one would like to have a comprehensive picture of the distribution of HG violations. However, we do not see a way to achieve this systematically for the entire lexicon. The primary obstacle is, we believe, an important property of the phenomenon itself: judgments about lexical entailment are inherently messy because of the flexible ways in which people refine meanings in context. As a result, there often isn't a single objective answer to the question of whether two disjuncts stand in an entailment relation. For instance, whereas the disjuncts in (6a) have a dependable semantic relationship, (6b) is much less clear-cut.

(6) a. "The nuptials will take place in either France or Paris."

b. "In 1940, 37 percent of us had gone to a church or synagogue in the last week."

Some speakers have firm judgments that *church* and *synagogue* exclude each other, making (6b) clearly HG-respecting. However, it is easy to find uses of the phrase "synagogues and other churches", which presuppose that a synagogue is a kind of church. And we should take care even with our assertion that *France* and *Paris* invariably stand in an entailment relation. In contexts where France is being construed in terms of its countryside, or Paris in terms of its particular urban charms, *France* could come to mean something more like 'Paris outside of France'. The important thing for our purposes is that the insight behind HG shines through this uncertainty: no matter what one's initial view of the lexicon is, a speaker's use of a disjunction *X or Y* raises the likelihood that the disjuncts are semantically disjoint in her currently preferred lexicon. The speaker will be perceived as endorsing such an opinionated view of the language, at least for the current conversation, and the listener can either adopt that assumption or push back.

This LEXICAL UNCERTAINTY motivates our own explanation for why speakers utter HG-violating disjunctions. In broad terms, we say that such examples convey that the speaker is treating the two terms as exclusive. There are many potential motivations for this. Perhaps the most mundane is that the speaker simply lexicalizes the two terms as exclusive. The disjunction is likely to be easily justified in such cases, as it might be the most efficient and direct way of identifying the semantic union of the two terms.

More interesting are cases in which the speaker's disjunction seems to be part of an attempt to manage the listener's inferences. For instance, a speaker who uses the phrase *cheap or free* to describe an object or service might be concerned that using *cheap* alone will trigger

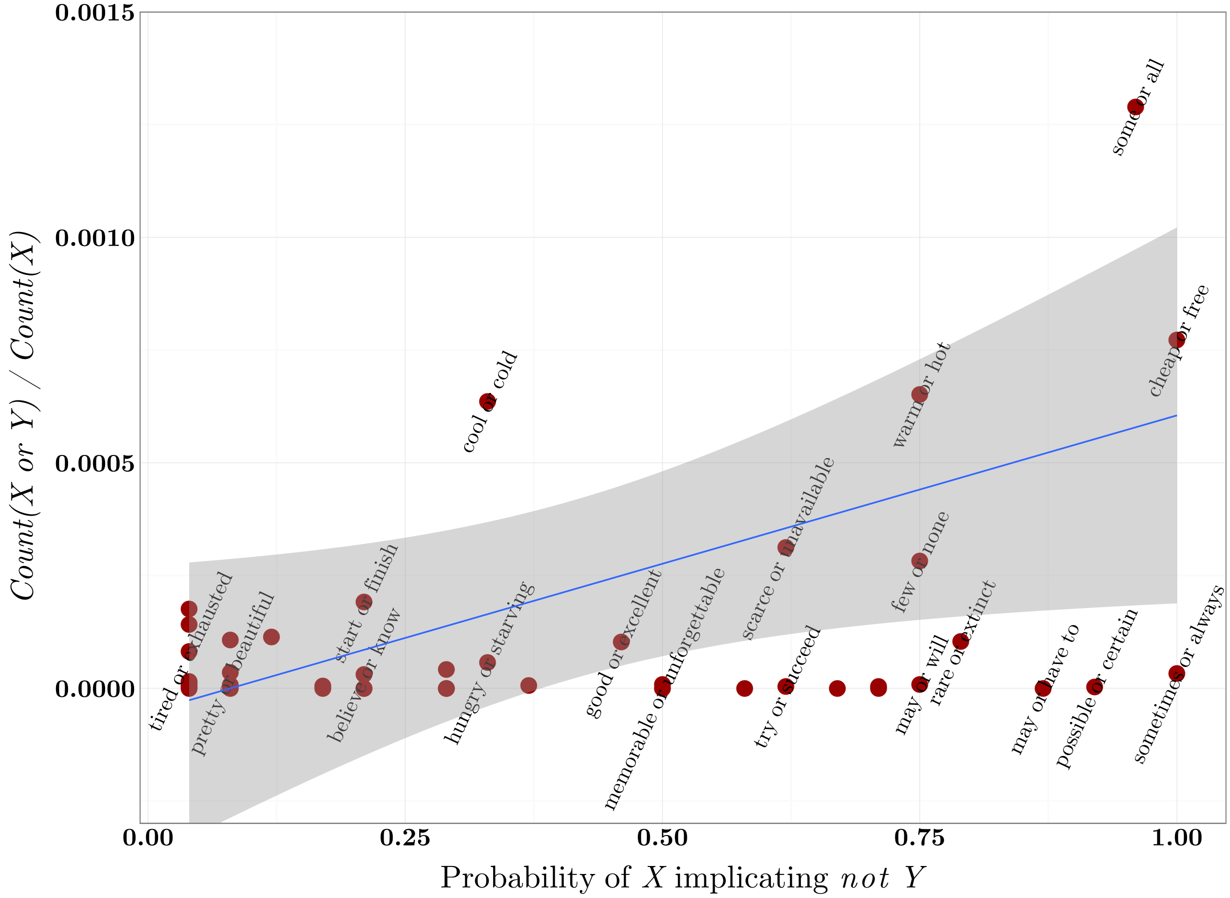


Figure 1: The relative frequency of $X\,or\,Y$ is predicted by the probability of X implicating *not* Y. All the data points are given (red dots) and included in the analysis. For the sake of readability, only a subset of the associated disjunctions are printed.

a scalar implicature (Hirschberg 1985) excluding the possibility that the object or service is free. The HG violation then serves to cancel or block this unwanted inference. Chemla (2013) studies this class of inferences, presenting suggestive evidence that the frequency of disjunctions $X\,or\,Y$ (X entailed by Y) is positively correlated with the likelihood that X conversationally implicates *not* Y as estimated by the experimental results of van Tiel et al. (2014). This connection is anticipated by Hurford's (1974) own analysis of disjunctions of scalar terms that seem to violate his generalization, and Chemla's findings suggest that it holds for a broad range of such cases.

Chemla's experiment relies on the hit counts in Google search results, which are notoriously unreliable (Liberman 2005), so we reproduced his main finding using the Google Books data set (Michel et al. 2011), pooling all the English-language tables and restricting attention to books from 1960 or later to avoid the encoding difficulties that plague earlier texts in that corpus. We also use a slightly more direct method than Chemla: we fit a simple linear regression in which the probability of X implicating Y is used to predict the relative frequency of $X\,or\,Y$. The implicature probabilities are from van Tiel et al.'s (2014) results, and the relative frequencies are given by $\mathit{Count}(X\,or\,Y)/\mathit{Count}(X)$, where $\mathit{Count}(\varphi)$ is the token count for the word or phrase φ in our Google Books sub-corpus. Our rationale (defined by Chemla) is that the prevalence of the implicature will positively correlate with the frequency of the disjunction: all other things being equal, the more likely the implicature, the more need there will be to block it. And this is what we find; the linear regression is significant ($p = 0.04$), suggesting a systematic relationship. Figure 1 summarizes this experiment.

Blocking a potential scalar implicatures is just one of the motivations a speaker might have for uttering a disjunction that superficially violates HG. Blocking an I(nformativeness)-implicature (Levinson 2000) might also be a motivating factor. Generally speaking, an I-implicature involves some kind of reasoning to a prototypical case — e.g., saying *The cup is on the table* implicates that the cup is in direct contact with the table. HG-violating disjunctions could be motivated by the possibility that a general term on its own would be understood as referring only more narrowly to a set of salient sub-kinds. For instance, at a busy marina in water-skiing country, *boat* might come to identify just motorboats. In that context, a speaker wishing to state a rule or regulation about all watercraft might use *boat or canoe* or *boat or kayak* to ensure that these non-motorized cases are included, lest people assume (or feel licensed to act as if they can assume) that these rarer kinds of boat are exempt. Such inferences often resemble scalar inferences, but a generalized notion of prior likelihood steers the calculation to a specific sub-kind among potentially many that might be at the same semantic or conceptual level.

In these implicature-blocking scenarios, the speaker is concerned that the general term X will be construed as $[\![X]\!] - p$ for some p overlapping with $[\![X]\!]$, and the disjunction hedges against that possibility. This is a DEFENSIVE position; the speaker's own lexicon might allow her to use just the general term to convey her intentions, but she is concerned that the listener will adopt a more restrictive interpretation. The costs of disjunction are therefore worth paying even if the disjunction adds no new information given the speaker's lexicon. However, the speaker can play a more active role as well, using disjunctions to instruct the listener about the correct lexicon. Our Section 1 example containing *angioplasty or surgery* seems to be an instance of this: the disjunction conveys secondary information that *angioplasty* and *surgery* will be treated as separate options in the discourse (Simons 2001). If HG were adopted as an explicit theoretical constraint, then the possibility of doing this would more or less follow. We would just require the additional premise that the listener is charitable and so will try to find an acceptable construal of the utterance. The model we develop in Section 3 also supports this reasoning, but it has the advantage of requiring no independent statement of HG.

2.2 Definition and Identification

Disjunctions like *wine lover or oenophile* seem to fly in the face of the Hurfordian pressure reviewed just above. Rather than avoiding overlap, they seem to embrace it, conveying something approximating identity.

Definitional disjunctions are even more contextually restricted than HG-violating ones, so identifying generalizations regarding their usage conditions is difficult. The examples in (7) were obtained by annotating the disjunctive clauses in a sample of 98 TED talks. These talks seem ideal for finding definitional uses, because the speakers are experts with broadly pedagogical aims. In addition, videos of the talks are available online, which made it possible for us to use intonation and other cues to try to verify that the speakers had definitional intentions. These seven examples were drawn from a set of 344 disjunctive clauses, for a rate of only about 2%, despite the context being conducive to such uses. The examples are given with the punctuation from the transcripts that TED provides.

(7) a. "...more disorder, or 'entropy,' over their lifetimes"
(Alex Wissner Gross, 'A new equation for intelligence', 5:51)

b. "...by the gametes, or the sperm and the eggs."
(Carin Bondar, 'The birds and the bees are just the beginning', 3:02)

c. "This is Toxoplasma gondii, or Toxo, for short, ..."
(Ed Yong, 'Zombie roaches and other parasite tales', 9:33)

d. "the Carnegie Airborne Observatory, or CAO"
(Greg Asner, 'Ecology from the air', 2:22)

e. "We call this project VocaliD [vəʊkælɪdi], or vocal I.D., ..."
(Rupal Patel, 'Synthetic voices, as unique fingerprints', 3:50)

f. "they can self-renew or make more of themselves ..."
(Siddharthan Chandran, 'Can the damaged brain repair itself?', 8:05)

g. "the endogenous stem cells, or precursor cells, ..."
(Siddharthan Chandran, 'Can the damaged brain repair itself?', 14:48)

As these examples suggest, speakers often (but not always) signal definitional intentions with ad hoc prosody, italics, quotation marks, and other devices, which points to the marked nature of the usage. However, it would be a mistake to dismiss definitional uses as an idiosyncrasy of English. These uses are widely attested in typologically diverse languages; we have examples from Chinese, German, Hebrew, Ilokano, Japanese, Russian, and Tagalog. Even languages that seem to have a dedicated 'definitional' or 'metalinguistic *or*' (e.g., Finnish, Italian) seem also to allow the regular *or* to play this role.

For our purposes, the most important property of these uses is that they convey a meaning that is secondary to the main content of the utterance — an extreme instance of a meaning that is not at-issue (Tonhauser et al. 2013; Dillon et al. 2014). This contrasts with overt definitions like *'oenophile' means 'wine-lover'* or *oenophile: wine-lover* (in a dictionary context). We think it is no accident that another strategy for conveying definitional information in this non-asserted, taken-for-granted manner is via apposition, as in *oenophile ('wine-lover')*, since appositives too are often recruited to convey secondary, supporting information (Potts 2005, 2012; Syrett & Koev 2014). In this respect, the relevant inference resembles the exclusivization pressure identified by HG: both seem to emerge as side-effects rather than normal outputs. In our model (Section 3), both are in turn characterized as "meta-linguistic" — inferences about the lexicon rather than about the state of the world.

In addition, as with disjunct exclusivization, the relevant lexical inference might be temporary. For instance, in cases like *Internet or computer network*, the second phrase seems to be used as a rough-and-ready way of helping the listener bootstrap towards an understanding of what the Internet is. Even our wine-lover example involves only approximate synonymy; according to our intuitions, *wine lover or oenophile* seems apt in a context in which the speaker wishes to use *oenophile* to elevate the concept to something more specific (or pretentious) than *wine lover* picks out. Similarly, the book title *A Geological History of Manhattan or New York Island* identifies Manhattan with New York Island while at the same time acknowledging the different histories and connotations of the two disjunct terms.

The speaker's motivations for using definitional disjunction are varied. Such readings seem to arise most easily when the speaker is mutually and publicly known to have EXPERTISE in the domain covered by the terms and the listener is mutually and publicly known to be inexpert in that area. In such cases, the speaker can use the disjunction to convey information about her preferred lexicon, fairly certain that the listener will be receptive. This characterization feels broadly fitting for our TED examples in (7). Some uses in this area have a self-corrective feel; the speaker, simulating her listener, might realize that she used a term that was inappropriate somehow — obscure, inexact, impolite (Clark 1994; Clark & Krych 2004). Disjunction provides a means for making amends fairly seamlessly. The following examples illustrate this strategy; in both, the speaker seems to realize that the first disjunct is potentially obscure and so rescues the phrase via definitional disjunction.

(8) a. "and e-waste is reported, or electronic waste is reported, by the UN, . . . "
(Leyla Acarogluk, 'Paper beats plastic', 14:10)

b. "Combined, the APR or annual percentage rate can be astronomical"
(Lou Dobbs Tonight, October 4, 2005)

While speaker expertise seems to be a genuine prerequisite, the listener's knowledge seems to impose little on felicitous uses (though it is certainly relevant). We find natural uses of this strategy when there is no direct information about the listener, but rather just a general assumption that one of the terms is relatively unknown. For instance, a newspaper article might contain *wine lover or oenophile* without presuming that all its readers are ignorant of the term; rather, such a use would seem to presuppose only that *oenophile* is relatively unknown, or obscure enough that it's useful to reinvoke its definition.

At the other end of the spectrum, the listener might actually be presumed to know the term, but the speaker sees social value in conveying that she shares this view. This could be because the speaker would like to display expertise, as when an ambitious pupil seeks to convey competence to a teacher. Similarly, a speaker of Australian English might use a phrase like *lift or elevator* with an American colleague to signal a willingness (or ironic lack thereof) to use the American form *elevator*.[2] And definitional uses also arise when the speaker and listener are both experts in the domain and see value (jointly or just in the current speaker's eyes) in using a word in a specialized sense in order to name a concept efficiently (e.g., the hypothetical academic paper title *What motivates the snobbish wine lover or 'oenophile' and how does he differ from the casual drinker?*).

For these reasons, we propose the following, more general set of desiderata that we believe must hold in order for definitional interpretations to be licensed. First, the discourse participants must have a mutual interest in communicating not only about the world but also about their language and arriving at a refined — even if context-specific and fleeting — joint understanding of it. Second, the discourse participants should share as a background assumption that the speaker and listener are willing to coordinate on the lexicon that the speaker seems to be using. That is, there must be tacit agreement between speaker and

[2]Our thanks to James Collins for bringing such uses to our attention.

listener that utterance interpretation can proceed on the assumption of speaker expertise in the language of the domain of the conversation. Third and finally, the cost (in the Gricean sense) of using a disjunction must be fairly small, all things considered, else it is hard to see how the speaker could justify using a disjunction $X\,or\,Y$ to convey simply $[\![X]\!]$. That is, whatever the costs of using the verbose form, they must be worth paying in virtue of the benefits of identifying (for the purposes of the current talk exchange) the meanings of the two disjuncts.

3 Modeling Communication under Conditions of Speaker Expertise

We now describe our model of pragmatic reasoning. Our presentation is somewhat compact. Readers wishing to more fully explore the model are referred to the website for this paper, which provides implementations of this model (as well as those of Frank & Goodman 2012, Bergen et al. 2014, and Smith et al. 2013) and includes code for calculating all of the examples we review here.

In our model, production and interpretation are based in a recursive process in which speaker and listener agents reason about each other reasoning about each other. At the lowest levels of our model, these agents communicate using a single lexicon. However, we do not actually assume that a single lexicon is mutually and publicly recognized as the set of core conventions of the language. Rather, our model aggregates over many possible lexica, thereby allowing the agents to negotiate the structure of the language itself even as they communicate. Figure 2 summarizes this picture schematically; this section is devoted to explicating these agents and their relationships.

The core structures of our model are given in (9). Intuitively, we imagine that a speaker and listener are playing a game in which the speaker privately observes a state $w \in W$ and produces a message $m \in M$ on that basis, given the context defined by the signaling system. The listener then uses m to guess a state $w' \in W$. The communication is successful just in case $w = w'$. The agents that we define are rational in the sense that, by reasoning recursively about each other's behaviors, they can increase their chances of success at this signaling game.

(9) a. W is a set of states (worlds, referents, propositions, etc.).
b. M is a set of messages containing a designated 'null' message $\mathbf{0}$.
c. $\mathcal{L}^* : M \mapsto \wp(W)$ is a semantic interpretation function. $\mathcal{L}'(\mathbf{0}) = W$.
d. $P : \wp(W) \mapsto [0, 1]$ is a prior probability distribution over states.
e. $C : M \mapsto \mathbb{R}$ is a cost function on messages.

Clause (9b) designates a null message $\mathbf{0}$ and clause (9c) includes a stipulation that $\mathbf{0}$ is true in all states in all lexica. It can be thought of as a catch-all for the numerous messages in the language that are not discriminating in the context defined by the signaling system. No matter how big and complex the examples, such a message is always justified. Introducing $\mathbf{0}$ also helps ensure that various calculations in the model are well-defined. (For alternative methods of ensuring definedness, see Jäger 2012 and Bergen et al. 2014.)

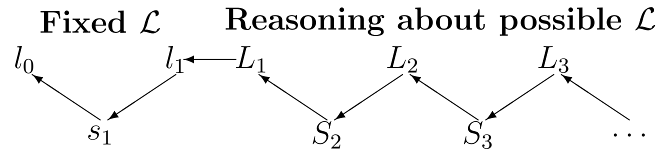


Figure 2: Summary of model structure.

3.1 Simple State/Message Signaling

As Figure 2 shows, our model defines an intuitive hierarchy of agents. The most basic are l_0, s_1, and l_1. In intuitive terms, these agents make meaning in a Gricean fashion: the pragmatic listener l_1 reasons about the pragmatic speaker s_1 reasoning about the 'literal' (purely semantic) listener l_0. These agents reason in terms of a fixed lexicon $\mathcal{L}$, which, for our purposes, can be thought of as a standard semantic interpretation function, as in (9c). These agents suffice to define a version of the Rational Speech Acts model of Frank & Goodman (2012) and Goodman & Stuhlmüller (2013).

The starting point is the literal listener, l_0. This agent simply turns $\mathcal{L}$ into a probabilistic formulation that can be used for decision-making in the presence of communicative indeterminacy. In short, given message m, this agent assigns 0 probability to states in which m is false. For the rest, its estimated probability for each state w is proportional to the prior $P(w)$.[3]

(10) $l_0(w \mid m, \mathcal{L}) \propto P(w)$ if $w \in \mathcal{L}(m)$, else 0

This literal listener is pragmatic only insofar as it incorporates the contextual prior P into (a distribution derived from) the truth conditions. Richer pragmatic inferences start to emerge as soon as we envision a speaker that can reason in terms of this listener and plan its utterances accordingly. The minimal such speaker is s_1, which is defined in terms of l_0:

(11) $s_1(m \mid w, \mathcal{L}) \propto \exp\left(\alpha \log\left(l_0(w \mid m, \mathcal{L})\right) - C(m)\right)$

This agent observes a world state w and, given $\mathcal{L}$, chooses a message m to try to communicate w. The definition appears cumbersome in its use of log and exp. These transformations are needed to ensure that all values are positive even with real-valued costs. At its heart, though, this agent is parallel to l_0 in that it combines a conditional distribution and a piece of contextual information — now costs on messages. The result is a distribution that one can imagine serving as the basis for decision-making in production: given that the agent would like to convey that a certain state holds, which message will do that most effectively for the l_0 listener? The real-valued parameter α controls the degree to which the speaker tries to capitalize on the distinctions l_0 encodes.

The first pragmatic listener is l_1. It is parallel to l_0 except that it reasons, not in terms of the original lexicon, but rather in terms of s_1 reasoning about l_0 reasoning about the lexicon. This agent is essentially a derived, pragmatically-enriched probabilistic interpretation function:

(12) $l_1(w \mid m, \mathcal{L}) \propto s_1(m \mid w, \mathcal{L})P(w)$

[3] '$P(a \mid b) \propto X$' is read 'the value $P(a \mid b)$ is proportional to the value X'. To obtain normalized probabilities, one divides each value X by the sum of all the values X' obtained by replacing a by one of its a'.

$\mathcal{L}^*$	w_1	w_2	w_3
p	T	T	F
q	T	F	T
$p\,\&\,q$	T	F	F
p *or* q	T	T	T
0	T	T	T

$\longleftarrow$

l_0	w_1	w_2	w_3
p	.5	.5	0
q	.5	0	.5
$p\,\&\,q$	1	0	0
p *or* q	.33	.33	.33
0	.33	.33	.33

l_1	w_1	w_2	w_3
p	.3	.7	0
q	.3	0	.7
$p\,\&\,q$	1	0	0
p *or* q	.17	.41	.41
0	.17	.41	.41

$\nwarrow$ $\swarrow$

s_1	p	q	$p\,\&\,q$	p *or* q	**0**
w_1	.33	.33	.25	.08	0
w_2	.8	0	0	.2	0
w_3	0	.8	0	.2	0

Figure 3: Simple state/message signaling with disjunction. $P(w_i) = 1/3$ for all states $w_i \in W$, $C(or) = C(and) = 1$, and $\alpha = 1$. Listener l_1's best inferences are in gray. The recursive process separates disjunction and conjunction, and it also separates disjunction from each of its disjuncts.

Figure 3 shows how this model derives basic scalar implicatures. We've used disjunction to illustrate, but the reasoning holds wherever general terms compete with more specific (entailing or merely overlapping) ones. In terms of the communication game, suppose the speaker produced the message $p\ or\ q$. For the literal listener l_0, all three worlds have equal probability given this message, making it unlikely that the two agents will succeed in communication. However, the first pragmatic listener l_1, reasoning in terms of s_1, has a greater chance of success. It has already learned to separate the related terms: $p\ or\ q$ conveys that $p\,\&\,q$ is false, and the atomic propositions convey biases not present in their disjunction. One can also see that the speaker seeks to avoid ambiguities. For instance, where a unique atomic proposition is true, the speaker opts for it, thereby creating less uncertainty in the listener than a disjunction would.

This basic model has been shown to achieve good quantitative fits to experimental results (Degen & Franke 2012; Stiller et al. 2011) and to contribute to artificial agents that communicate effectively with each other to solve a collaborative task (Vogel et al. 2013). One can also generalize l_1 and s_1 to allow them to recursively respond to each other, which strengthens the scalar inferences seen in Figure 3. In our model, there is no further recursion of these lexicon-specific agents, but we do allow further recursion of the agents we define next.

3.2 Reasoning under Lexical Uncertainty

In the model defined by l_0, s_1, and l_1, the agents condition on (take as given) a fixed, shared lexicon. However, the data in Section 2 show that the lexicon is not known precisely, but rather negotiated during interactions. We now bring that insight into our model using the lexical uncertainty technique first introduced by Bergen et al. (2012, 2014). The first step is to define a space of lexica $\mathbf{L}$ and a probability distribution over them $P_{\mathbf{L}}$:

(13) a. $\mathbf{L} = \{\mathcal{L}' : \mathcal{L}'(\mathbf{0}) = W \,\&\, \forall m \in M - \{\mathbf{0}\}, \mathcal{L}'(m) \neq \emptyset \,\&\, \mathcal{L}'(m) \subseteq \mathcal{L}^*(m)\}$

b. $P_{\mathbf{L}} : \wp(\mathbf{L}) \mapsto [0, 1]$ is a prior probability distribution over lexica.

Clause (13a) defines a complete set of refinements of a basic lexicon: every lexical item can denote a subset of the space it denotes in the base lexicon $\mathcal{L}^*$, and all combinations of these refinements are available. The only requirements are that messages always have non-empty denotations and that the null message **0** always denotes the full set of states. We should emphasize, though, that these assumptions are not essential to our model. In the extreme case, we could admit every lexicon derivable from the messages M and states W, and then use the lexicon prior $P_{\mathbf{L}}$ to provide some structure to the space — say, by assigning low but non-zero probability to lexica that are not strict refinements, and zero probability to lexica in which **0** is not universal or some messages have empty denotations. This is a strictly more general perspective than the one given by (13). We do not explore these options further in this paper, but see Kao et al. 2014a,b for discussion of non-refinement pragmatic enrichment involving non-literal language.

Our first 'lexical uncertainty' agent is L_1, which is based on the model of Smith et al. (2013). Given a message, this agent makes joint lexicon–state inferences over the space $W \times \mathbf{L}$:

(14) $L_k(w, \mathcal{L} \mid m) = L_k(w \mid m, \mathcal{L}) L_k(\mathcal{L} \mid m)$

$L_k(\mathcal{L} \mid m) \propto P(\mathcal{L}) \sum_{w \in W} S_k(m \mid w, \mathcal{L}) P(w)$

This agent is defined for levels $k \geqslant 1$ as long as we assume that $S_1 = s_1$ (higher levels use the speaker in (15)) and $L_1(w \mid m, \mathcal{L}) = l_1(w \mid m, \mathcal{L})$. The first term uses a fixed-lexicon inference and the second, $L_k(\mathcal{L} \mid m)$, encodes the extent to which the current message biases in favor of the lexicon $\mathcal{L}$.

Our 'expertise' speaker responds to this lexical uncertainty listener. We assume that this speaker does have a specific lexicon in mind (this is the sense in which it is expert). More precisely, this speaker is taken to observe state–lexicon pairs and produce messages on that basis. It is defined for all $k > 1$:

(15) $S_k(m \mid w, \mathcal{L}) \propto \exp\left(\alpha \log\left(L_{k-1}(w \mid m, \mathcal{L})\right) + \beta \log\left(L_{k-1}(\mathcal{L} \mid m)\right) - C(m)\right)$

In broad terms, this agent is similar to s_1 except that it includes a new term $L_{k-1}(\mathcal{L} \mid m)$ encoding the information each message conveys about the lexicon. The importance of this information is controlled by a new real-valued parameter β. The relative weights of α and β govern the relative importance of conveying information about the world and information about the lexicon. As α grows, the agents venture riskier Gricean behavior: the speaker tries to extract as many distinctions as possible from its estimate of the listener, and it chooses utterances on this (hyper-)rational basis. The pattern is similar for β, but the target becomes exchanging information about the language itself. If we set $\beta = 0$, then the agents places no value on communicating about the lexicon, and the model reduces to a variant of the lexical uncertainty model of Bergen et al. (2014). As β rises, communicating about the language itself becomes more important. The precise relationship between α and β is extremely complex; much of our investigation of disjunction focuses on understanding it.

Let's return to the schematic diagram in Figure 2. Intuitively, if one begins at, say, L_2, then the reasoning flows down through the uncertainty agents S_2 and L_1, at which point it splits apart into lower-level agents that reason about specific lexica. Alternatively, one can imagine the lexicon-specific inferences flowing up to be pooled together by L_1, which then makes joint inferences about the state and the lexicon.

Figure 4 seeks to show how the model works with a simple case involving a scalar inference. In the figure, the bottom row specifies the lexica defined by (13a) with $\mathcal{L}^*$ as the starting point; the more specific term has no further refinements, but the general term has two further refinements, resulting in three lexica. The lexicon-specific agents l_0, s_1, and l_1 reason about these lexica. At the L_1 level, we have given just the joint table for the message *cheap*, but there are similar tables for the messages *free* and **0**. Similarly, we've given only one conditional probability table for S_2; there are similar tables for $\mathcal{L}^*$ and $\mathcal{L}_1$. Finally, we depict the *cheap* table for L_2. For this example, we set $\alpha = 1$ and $\beta = 2$. The relatively large β parameter means that the speaker values communicating about the lexicon, and we can see the effects of this in L_2, which displays a dispreference for $\mathcal{L}_1$, equivalently, a preference for the lexica that distinguish *cheap* from *free*. More generally, even in this basic example, $L_2(\mathcal{L} \mid m)$ discriminates among lexica, meaning that this agent can use S_2's message to gain insights into its lexical preferences. If β is lowered, then this becomes less pronounced, and $\beta = 0$ means that $L_2(\mathcal{L} \mid m)$ becomes flat (treats all lexica as identical).

3.3 A Return to Simple Signaling

The full model can be unwieldy because of the lexicon inferences. One often wants to return to the intuitive picture of simple state/message signaling, as in examples like Figure 3. To achieve this, one can marginalize (or sum) over lexica, as in (16) and (17). Both provide useful summaries of the model's predictions.

(16) $L_k(w \mid m) = \sum_{\mathcal{L} \in \mathbf{L}} L_k(w, \mathcal{L} \mid m)$

(17) $S_k(m \mid w) \propto \sum_{\mathcal{L} \in \mathbf{L}} S_k(m \mid w, \mathcal{L}) P_{\mathbf{L}}(\mathcal{L})$

Figure 5 uses these equations to summarize the inferences for L_2 and S_2 from Figure 4. We no longer get insights into the agents' lexical preferences and inferences, but we do see that they are computing scalar inferences in the manner we saw for the simpler model in Figure 3.

4 Compositional Lexical Uncertainty

There are two shortcomings of the model as presented so far that are important to correct before studying Hurfordian and definitional usage patterns. Both of these shortcomings are evident in the treatment of disjunction in Figure 3.

First, if we took the lexicon given there as the base lexicon and applied lexical uncertainty to it using (13a), the results would be highly unintuitive because many lexica would fail to respect the desired semantic relationships that hold among the messages. For example, there would be lexica in which *p or q* strictly entailed *p*. The culprit here is that lexical uncertainty, applied naively, does not respect compositional semantics. To address this, we follow Bergen et al. (2014) and Levy et al. (To appear) in applying uncertainty only to the true lexical items, and then close each lexicon under the compositional operations of disjunction and

L_2

heard *cheap*	w_1	w_2
$\mathcal{L}^*$	.12	.29
$\mathcal{L}_1$	.04	.22
$\mathcal{L}_2$	.04	.29

↓

S_2

prefers $\mathcal{L}_2$	*cheap*	*free*	**0**
w_1	.15	.85	0
w_2	1	0	0

↓

L_1

heard *cheap*	w_1	w_2
$\mathcal{L}^*$	.12	.35
$\mathcal{L}_1$	.18	0
$\mathcal{L}_2$	0	.35

↙ ↓ ↘

l_1

	w_1	w_2
cheap	.25	.75
free	1	0
0	.25	.75

	w_1	w_2
cheap	1	0
free	1	0
0	0	1

	w_1	w_2
cheap	0	1
free	1	0
0	.5	.5

↓ ↓ ↓

s_1

	cheap	*free*	**0**
w_1	.33	.67	0
w_2	.99	0	.01

	cheap	*free*	**0**
w_1	.5	.5	0
w_2	0	0	1

	cheap	*free*	**0**
w_1	0	1	0
w_2	1	0	0

↓ ↓ ↓

l_0

	w_1	w_2
cheap	.5	.5
free	1	0
0	.5	.5

	w_1	w_2
cheap	1	0
free	1	0
0	.5	.5

	w_1	w_2
cheap	0	1
free	1	0
0	.5	.5

↓ ↓ ↓

$\mathcal{L}$

$\mathcal{L}^*$	w_1	w_2
cheap	T	T
free	T	F
0	T	T

$\mathcal{L}_1$	w_1	w_2
cheap	T	F
free	T	F
0	T	T

$\mathcal{L}_2$	w_1	w_2
cheap	F	T
free	T	F
0	T	T

Figure 4: Illustration of the full model in action. The priors over worlds and lexica are all flat, message costs are all 0, $\alpha = 1$, and $\beta = 2$. The high β means that L_2 shows a notable dispreference for the undiscriminating lexicon $\mathcal{L}_1$.

	w_1	w_2
cheap	.2	.8
free	1	0
0	0	1

(a) L_2

	cheap	*free*	**0**
w_1	.23	.77	0
w_2	.92	0	.08

(b) S_2

Figure 5: Return to simple signaling for the simple scalars example in Figure 4.

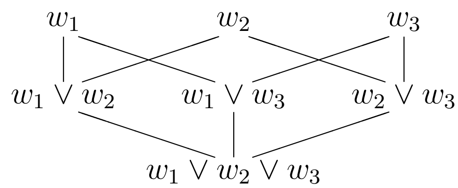


Figure 6: State space with disjunctive closure.

conjunction. Thus, in Figure 3, only p and q can be refined. This results in nine distinct lexica. Each of these is then expanded to include $p\,\&\,q$ and $p\; or\, q$, with their meanings determined by whatever meanings the atomic expressions have.

Second, the world space in Figure 3 does not create any room for the speaker to be uncertain. The pretense is that the speaker always observes the state perfectly (up to the level of granularity being measured) and then seeks to communicate that observation. This gives somewhat unintuitive results for disjunction. Disjunction is naturally used to convey speaker uncertainty, but we currently allow no room for such uncertainty. To address this, we close the state space under joins. For the set of atomic states $\{w_1, w_2, w_3\}$, this results in the semilattice in Figure 6. Interpretation via $\mathcal{L}$ then proceeds as one would expect: disjunction corresponds to union and conjunction to intersection. For example, if $\mathcal{L}(p) = \{w_1, w_2\}$ and $\mathcal{L}(q) = \{w_1, w_3\}$, then $\mathcal{L}(p\; or\, q) = \{w_1, w_2, w_3\}$. After join closure, this denotes the entire state space represented in Figure 6. Conversely, $p\,\&\,q$ denotes $\{w_1\}$, which does not change under closure. (We could expand the space of meanings to include meets, or even to the full Smyth powerlattice with states like $w_1 \vee (w_2 \wedge w_3)$, as in Levy & Pollard 2001, but we stick with the join space to keep the presentation as simple as possible.)

Formally, the model we have developed here is an instance of the compositional lexical-uncertainty model of Bergen et al. (2014) and Levy et al. (To appear), with two elaborations: we have placed value on transmission of the lexicon from speaker to listener (equation 15) and specified the structure of the state space for the classes of utterances presently under consideration. Figure 7 summarizes how our model behaves in this setting. We begin from a base lexicon $\mathcal{L}^*$ in which $\mathcal{L}^*(p) = \{w_1, w_2\}$ and $\mathcal{L}^*(q) = \{w_1, w_3\}$. We apply lexical uncertainty to obtain nine distinct lexica, close their state space under joins, and close their messages under disjunction and conjunction. We run the full model up to level L_2, and then marginalize L_2 and S_2 as in (16) and (17), respectively. The results are closely aligned with the comparable tables in Figure 3, in that we obtain all of the interpretive scalar implicatures and predict largely the same speaker behavior. The one major improvement is that we now directly capture the intuition that disjunction can signal speaker uncertainty: the speaker's best choice for state $w_2 \vee w_3$ is $p\; or\, q$.

5 Analysis

We now show how our model can be used to characterize a range of Hurfordian and definitional capacities of disjunction. We first illustrate the effects using parameters chosen by hand, giving an example of definitional inference in Section 5.1 and an example of Hurfordian inference in Section 5.2. We then show in Section 5.3 how our model derives the "defensive-speaker" Hurfordian disjunctions discussed in Section 2.1. Finally, in Section 5.4, we explore the space of parameters more fully to try to determine which settings — which

	w_1	w_2	w_3	$w_1 \vee w_2$	$w_1 \vee w_3$	$w_2 \vee w_3$	$w_1 \vee w_2 \vee w_3$
p	.25	.51	0	.24	0	0	0
q	.25	0	.51	0	.24	0	0
$p \,\&\, q$	1	0	0	0	0	0	0
$p\ or\ q$	.04	.06	.06	.18	.18	.28	.21
0	0	.12	.12	.16	.16	.2	.24

(a) L_2

	p	q	$p \,\&\, q$	$p\ or\ q$	**0**
w_1	.29	.29	.4	.02	0
w_2	.96	0	0	.04	0
w_3	0	.96	0	.04	0
$w_1 \vee w_2$	.79	0	0	.2	.01
$w_1 \vee w_3$	0	.79	0	.2	.01
$w_2 \vee w_3$	0	0	0	.95	.05
$w_1 \vee w_2 \vee w_3$	0	0	0	.93	.07

(b) S_2

Figure 7: Simple signaling for compositional disjunction.

approximations of the context — deliver each type of reading.

Throughout our illustrations, we assume that X is the target for refinement, that is, the general term for Hurfordian inferences and the unknown term for definitional ones. From the listener's perspective, we are concerned to see when $A\ or\ X$ gives rise to the inference that $[\![A]\!] \cap [\![X]\!] = \emptyset$ (Hurfordian reading) and when $A\ or\ X$ gives rise to the inference that $[\![A]\!] = [\![X]\!]$ (definitional). (From now on, we use $[\![m]\!]$ as a shorthand for 'the listener's construal of the message m'.)

It's worth highlighting again the dynamics of the model that facilitate these uses. First, we assume that the speaker has a preferred lexicon, and that she will choose her messages in part to help convey this preference. Of course, the recursive nature of the model means that the listener (or, rather, the speaker's expectations about the listener) are involved in this planning. Second, we assume that the listener is at least somewhat willing to defer to the speaker with regard to the best lexicon to use in context. Again, this deference is mediated by the recursive nature of the model, which defines all aspects of communication as a (boundedly) rational collaboration. We take these basic assumptions to be necessary, but not sufficient, for communicating in language about language.

In our model, the importance of communicating about the world is governed by the parameter α, and the importance of lexical communication is governed by the parameter β. If β is set to 0, then communication about the lexicon has no intrinsic value to the conversational agents. Where β is positive, the ratio of α to β is a rough guide to the relative importance of communicating about the world and communicating about the lexicon (though message costs and other facts about the context also play a role).

5.1 Definitional Contexts

We begin with definitional readings because they more clearly motivate the structure of our model. The central question is when $A\ or\ X$ is intended and construed as equivalent to $[\![A]\!]$. From the listener's perspective, this involves an inference that (perhaps roughly speaking, and perhaps just for the current context) $[\![A]\!] = [\![X]\!]$, where X is the unknown term. From the perspective of production, we want to know when a speaker will favor producing $A\ or\ X$ given that she favors a lexicon in which $[\![A]\!] = [\![X]\!]$ and observes a state equivalent to the literal meaning of $[\![A]\!]$. The full characterization of this phenomenon involves many complex factors relating to social status, mutual, public knowledge about knowledge, and so forth. We saw in Section 2 that the motivations are diverse and subtle. Our goal is to home in on the core aspects of these inferences.

Figure 8 summarizes a basic context in which the uncertainty listener makes a definitional inference and the speaker inclines toward definitional intentions. The context has three atomic states and three basic messages. The initial lexicon $\mathcal{L}^*$ gives rise to three refinements, pictured in the bottom row of the figure. (To save space, we depict just three of the messages.) Crucially, β is larger than α, encoding the primacy of lexical information over world information (though both remain highly relevant), and the cost of disjunction is low: 0.01.

The progression from bottom to top in Figure 8 helps explain how the model arrives at definitional construals and intentions. At L_1, the listener is undecided about whether the disjunction is definitional or Hurfordian. The information contained in the lexicon-specific reasoning patterns is enough to create a bias away from $\mathcal{L}^*$, but it is not enough to break the symmetry between the remaining inferences. At S_2, the effects of the high β parameter are beginning to be felt. Although this speaker has a bias away from using $A\ or\ X$ to convey a definitional intention, this bias is only slight, because of the low cost of disjunction. Additionally, the likelihood of S_2 saying $A\ or\ X$ is maximal when the speaker observes $\langle \mathcal{L}_2, w_1 \rangle$. Together, these factors strongly push L_2 to adopt a definitional construal of $A\ or\ X$. This construal is clearer (in the sense that the L_2 probability on $\langle \mathcal{L}_2, w_1 \rangle$ is higher) for $A\ or\ X$ than for any other utterance. Crucially, L_2 inferences for X (not shown for reasons of space) are split between $\langle \mathcal{L}_2, w_1 \rangle$ and $\langle \mathcal{L}_1, w_2 \rangle$ because both these world–lexicon pairs give high likelihood to S_2 uttering X. As a result, S_3, who plans her utterances based on the expected inferences of L_2, strongly prefers to use the disjunction when in a definitional state. Once this definitional interpretation is reached, it remains stable through further iterations of listener–speaker recursive reasoning (also not shown for reasons of space).

The lexicon used in Figure 8 is tiny, and one might wonder whether the definitional inference comes about via a spurious process of elimination, given that there are only three atomic messages and one has a very general meaning. To address this concern, we ran the same simulation with a larger lexicon: five atomic messages and four atomic states. This is of course still unrealistically small for a natural language lexicon, but it certainly provides a much larger space for potential lexical refinements. The definitional reading arises in this lexicon with the same parameters as in Figure 8. This is unwieldy to visualize (there are fourteen distinct lexica), but the important thing is that the best lexicon is a definitional one. Here's the best lexical inference given the message $A\ or\ X$ from the speaker:

S_3

	A	B	X	$A or X$	$B or X$
observed $\langle \mathcal{L}_2, w_1 \rangle$	0	0	0	1	0
observed $\langle \mathcal{L}_1, w_2 \rangle$	0	0	0	0	1

↓

L_2

heard $A or X$	w_1	w_2	$w_1 \vee w_2$
$\mathcal{L}^*[A \mapsto \{w_1\}, B \mapsto \{w_2\}, X \mapsto \{w_1, w_2\}]$	0	0	.08
$\mathcal{L}_1[A \mapsto \{w_1\}, B \mapsto \{w_2\}, X \mapsto \{w_2\}]$	.01	0	.08
$\mathcal{L}_2[A \mapsto \{\mathbf{w_1}\}, B \mapsto \{w_2\}, X \mapsto \{\mathbf{w_1}\}]$	.77	0	.06

↓

S_2

	A	B	X	$A or X$	$B or X$
observed $\langle \mathcal{L}_2, w_1 \rangle$	.07	0	.48	.45	0
observed $\langle \mathcal{L}_1, w_2 \rangle$	0	.07	.48	0	.45

↓

L_1

heard $A or X$	w_1	w_2	$w_1 \vee w_2$
$\mathcal{L}^*[A \mapsto \{w_1\}, B \mapsto \{w_2\}, X \mapsto \{w_1, w_2\}]$	0	0	.23
$\mathcal{L}_1[A \mapsto \{w_1\}, B \mapsto \{w_2\}, X \mapsto \{w_2\}]$	0	0	.38
$\mathcal{L}_2[A \mapsto \{\mathbf{w_1}\}, B \mapsto \{w_2\}, X \mapsto \{\mathbf{w_1}\}]$	.38	0	0

↙ ↓ ↘

l_1

$\mathcal{L}^*$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
B	0	1	0
X	.02	.02	.96
$A or X$	.02	.02	.96
$B or X$	.02	.02	.96

$\mathcal{L}_1$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
B	0	1	0
X	0	1	0
$A or X$	.01	0	.99
$B or X$	0	1	0

$\mathcal{L}_2$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
B	0	1	0
X	1	0	0
$A or X$	1	0	0
$B or X$	0	.01	.99

↓ ↓ ↓

s_1

$\mathcal{L}^*$	A	B	X	$A or X$	$B or X$
w_1	.98	0	0	0	0
w_2	0	.98	0	0	0
$w_1 \vee w_2$	0	0	.2	.2	.2

$\mathcal{L}_1$	A	B	X	$A or X$	$B or X$
w_1	1	0	0	0	0
w_2	0	.33	.33	0	.33
$w_1 \vee w_2$	0	0	0	.33	0

$\mathcal{L}_2$	A	B	X	$A or X$	$B or X$
w_1	.33	0	.33	.33	0
w_2	0	1	0	0	0
$w_1 \vee w_2$	0	0	0	0	.33

↓ ↓ ↓

l_0

$\mathcal{L}^*$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
B	0	1	0
X	.33	.33	.33
$A or X$	.33	.33	.33
$B or X$	.33	.33	.33

$\mathcal{L}_1$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
B	0	1	0
X	0	1	0
$A or X$	.33	.33	.33
$B or X$	0	1	0

$\mathcal{L}_2$	w_1	w_2	$w_1 \vee w_2$
A	1	0	0
B	0	1	0
X	1	0	0
$A or X$	1	0	0
$B or X$	.33	.33	.33

Figure 8: Definitional intentions and construals, with only a subset of the states and messages displayed due to space constraints. Priors are flat. $\alpha = 5$; $\beta = 7$; $C(or) = 0.01$. The definitional construal becomes prominent at L_2, and definitional intentions emerge at S_3.

	L_2 hears $A\ or\ X$	w_1	w_2	$w_1 \vee w_2$
$\mathcal{L}_1$	$\left[A \mapsto \{w_1\},\, B \mapsto \{w_2\},\, X \mapsto \{w_2\}\right]$	0	0	.04
$\mathcal{L}_2$	$\left[A \mapsto \{\mathbf{w_1}\},\, B \mapsto \{w_2\},\, X \mapsto \{\mathbf{w_1}\}\right]$	.92	0	.03

Figure 9: Definitional inference for L_2 with X constrained to have an atomic meaning. Priors are flat. $\alpha = 5$; $\beta = 7$; $C(or) = 0.01$.

(18) $$\begin{bmatrix} A & \mapsto & \{w_1\} \\ B & \mapsto & \{w_2\} \\ C & \mapsto & \{w_3\} \\ D & \mapsto & \{w_4\} \\ X & \mapsto & \{w_1\} \end{bmatrix}$$

In our model, definitional uses are delicate in the sense that they arise only for a relatively narrow range of parameter settings. (We quantify this intuitive characterization in Section 5.4.) For example, if β is too close to α, or disjunction costs get too high, the listener fails to make the inference and the speaker tends to resort to using just A to convey w_1. This makes intuitive sense given the nature of these parameters. For example, if the speaker has little interest in communicating about the lexicon, or disjunction costs prohibit the wastefulness of using $A\ or\ X$ to convey $[\![A]\!]$, then the meanings disappear.

Nonetheless, it is worth asking whether there are ways in which the reading can be made more salient and robust. We have found that the rate of definitional intentions and perceptions is increased as we raise the prior expectation in favor of lexica in which the unknown term has a maximally general meaning; this corresponds intuitively to situations in which there is a mutual prior expectation that the word is unknown. In addition, we can further encourage pedagogical expectations by imposing the additional lexical requirement that the meaning of X be atomic (as atomic as the known terms in the lexicon). This would correspond to a situation in which it was clear that X was an alternative for some lexical item. Under these circumstances, the meaning of the known disjunct serves as a FOCAL POINT (Schelling 1960) that the speaker and listener can coordinate on for the meaning of the unknown word. Figure 9 is a minimal variant of (the topmost table in) Figure 8 in which this extra lexical requirement has been imposed. This removes $\mathcal{L}^*$ from consideration. The listener inference is numerically stronger, and it is paralleled by a similar increased probability on the speaker side for producing $A\ or\ X$ with definitional intentions. We will explore these atomic-meaning, focal-point contexts in more detail in Section 5.4.

5.2 Hurfordian Contexts

For Hurfordian readings, we seek to understand where and how $A\ or\ X$ gives rise to the inference that $[\![A]\!] \cap [\![X]\!] = \emptyset$, where X is again the unknown word. We assume that the overall meaning will be $[\![A]\!] \cup [\![X]\!]$, i.e., that the relevant information concerns just the lexicon.

Figure 10 presents a context in which L_2 arrives at this conclusion. The states and messages are as in our definitional example (Figure 8). Anticipating the meta-theory we develop in Section 5.4, we note that setting $\alpha > \beta$ is important for achieving this result. Intuitively,

L_2 hears $A \textit{ or } X$	w_1	w_2	$w_1 \vee w_2$
$\mathcal{L}^*\left[A \mapsto \{w_1\},\ B \mapsto \{w_2\},\ X \mapsto \{w_1, w_2\}\right]$	.02	0	.32
$\mathcal{L}_1\left[A \mapsto \{\mathbf{w_1}\},\ B \mapsto \{w_2\},\ X \mapsto \{\mathbf{w_2}\}\right]$	.04	0	.45
$\mathcal{L}_2\left[A \mapsto \{w_1\},\ B \mapsto \{w_2\},\ X \mapsto \{w_1\}\right]$	.03	0	.14

Figure 10: Hurfordian inference for L_2. Priors are flat. $\alpha = 2$; $\beta = 1$; $C(\textit{or}) = 1$.

this means that world information is valued more than lexical information. We also note that the cost of disjunction is relatively high. Where costs are high, the disjunction has to be justified. Letting the two terms overlap reduces the justification, whereas exclusivizing provides justification. In other words, the apparently undue prolixity of the disjunction (the more general term would seem to suffice!) supports the Hurfordian lexical inference.

Ours is a model of production as well as interpretation, so it's important to take the speaker's perspective as well. It's useful that the listener makes the desired inference upon hearing $A \textit{ or } X$, but the value of this observation is diminished if the speaker never uses this message to signal the relevant state–lexicon pair, especially since our corpus work (Section 2.1) suggests that speakers do in fact produce these utterances systematically. With the above parameters, our model predicts that they will produce such messages. Indeed, in this context, with these parameters, S_2's preferred message given observed state $w_1 \vee w_2$ and lexicon $\mathcal{L}_1$ from Figure 10 is $A \textit{ or } X$. This suggests that the strategy is a stable one in communication, in line with the intuitions we described in Section 2.1.

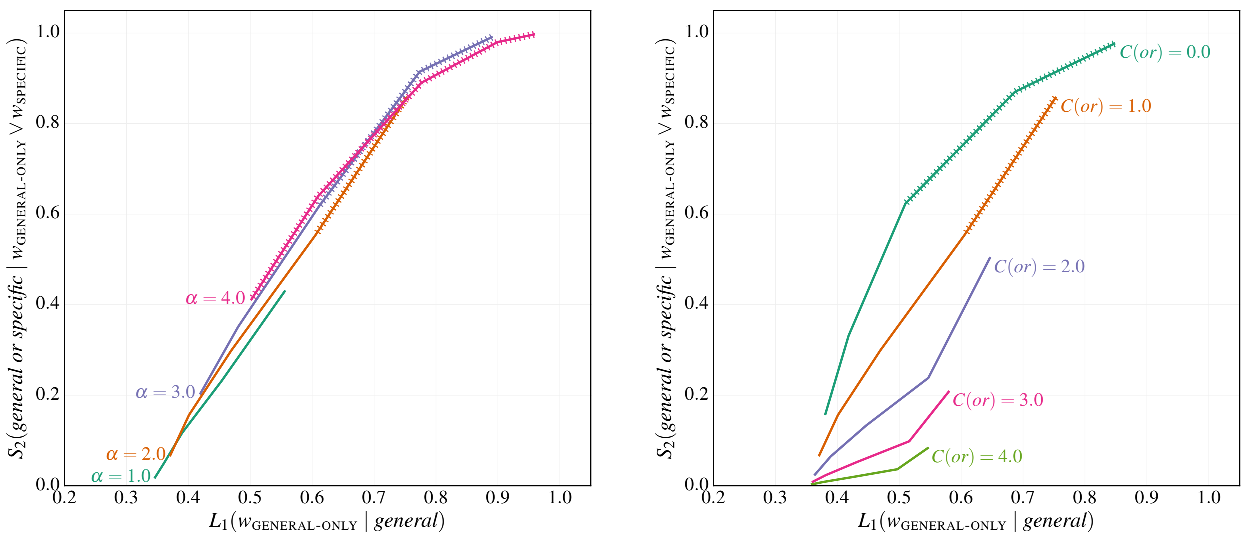
Once again, it is worth running the same simulation with a larger lexicon. We find that the above parameters deliver the same result with the lexicon size increased to five atomic messages and four atomic states: if the listener hears $A \textit{ or } X$ in this context, she infers that the speaker's lexicon is (19).

(19) $$\begin{bmatrix} A & \mapsto & \{w_1\} \\ B & \mapsto & \{w_2\} \\ C & \mapsto & \{w_3\} \\ D & \mapsto & \{w_4\} \\ X & \mapsto & \{w_2, w_3, w_4\} \end{bmatrix}$$

This is the 'minimal' Hurfordian lexicon: the listener doesn't infer anything about X except that it is disjoint from A. Once again, this inference is mirrored by a speaker preference for producing $A \textit{ or } X$ given observed state $w_1 \vee w_2$ and the lexicon (19).

5.3 Defensive Speakers: Q- and I-implicature Blocking

In Section 2.1, we identified two classes of potential implicature inferences that can drive speakers to produce HG violations: Q-implicatures, where using only the general term might create an upper-bounding inference, and I-implicatures, where using only the general term might result in inference to a specific, salient sub-kind. As we describe below, our model can derive speakers' use of HG violations to block both kinds of implicatures. It also sheds light on two aspects of these inferences: the circumstances that favor each type of implicature, and the nature of defensive implicature-blocking behavior on the part of the speaker.



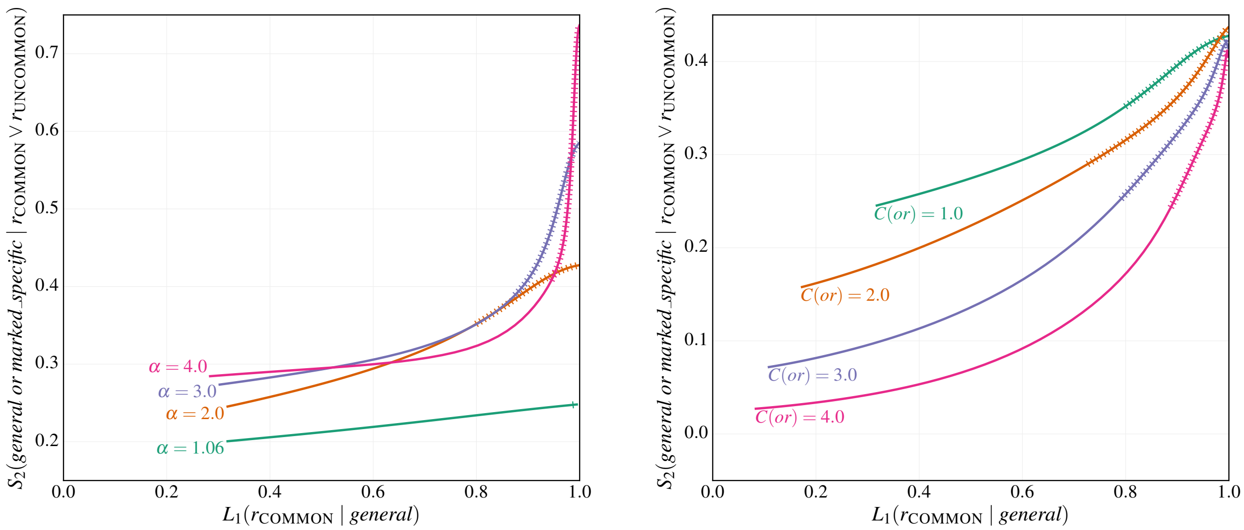
(a) Tracing five values of α with $C(or) = 1$.

(b) Tracing four values of $C(or)$ with $\alpha = 2$.

Figure 11: Q-implicature blocking. The variation in the x-axes is created by varying $C(\textit{specific})$ from 0 to 4, with lower costs corresponding to higher values (more implicature inferences from *general* to *not specific*). Dashes mark spans in which the speaker probability is maximal given the disjunctive observation, that is, where the HG violation is preferred.

Figure 11 reports on simulations involving the Q-implicature case. Intuitively, we have a general term like *cheap* and a specific term like *free*. The general term is true in both $w_{\text{GENERAL-ONLY}}$ and w_{SPECIFIC}, whereas the specific term is true only in w_{SPECIFIC}, creating an information state supporting a standard scalar implicature from the general term to the falsity of the specific one. In both panels of Figure 11, the x-axis tracks the probability that L_1, upon hearing the general term, draws an implicature inference that the specific term is false. We create this variation by changing the cost of the term *specific* from 0 to 4. As its cost rises, the listener's implicature becomes less probable. The y-axes depict the S_2 speaker's response to this listener upon observing the state $w_{\text{GENERAL-ONLY}} \vee w_{\text{SPECIFIC}}$. We are interested primarily in the conditions under which this speaker will produce the HG-violating disjunction *general or specific* having observed this state. The dashes on the lines mark spans in which an HG violation is the speaker's *preferred* choice here.

In both panels, the trend is the expected one, mirroring the corpus results from Figure 1. The patterns are heavily influenced by other parameters in the model, especially the setting of α and the cost of disjunction. For instance, consider the case where $\alpha = 2$, which is given by the orange line in Figure 11a. Here, when the listener probability reaches .6, the speaker's preferred move is to utter an HG-violating disjunction. With higher (more aggressive) α, the listener threshold drops to .5; at lower values of α, this particular speaker never violates HG. Similarly, in Figure 11b, where disjunction is free (green line), the speaker is willing to violate HG even when the listener's scalar inference is not all that strong; as the cost of disjunction goes up, the speaker needs a more compelling reason for the violation, that is, a stronger listener inference. If disjunction is sufficiently costly (e.g., the pink line), then this speaker never violates HG.



(a) Tracing four values of α with $C(or) = 1$. (b) Tracing four values of $C(or)$ with $\alpha = 2$.

Figure 12: I-implicature blocking. The variation in the x-axes is created by varying $P(r_{\text{COMMON}})$ from 0.33 to 0.99, with higher values for this prior corresponding to higher listener inferences (more implicatures from *general* to *specific_unmarked*). Dashes mark spans in which the speaker probability is maximal given the disjunctive observation, that is, where the HG violation is preferred.

The key mechanism for this Q-implicature example is variation in the cost of the specific term. To simulate I-implicatures, we instead vary the prior likelihood of the referents. Figure 12 summarizes the picture abstractly. Here, we have a general term like *boat* and two specific terms like *motorboat* and *canoe*, referring to objects r_{COMMON} and r_{UNCOMMON}, respectively. This would correspond to a situation where we work at a marina that is largely devoted to motorboats, with canoes infrequent and not especially salient. As in Figure 11, the x-axes track the inferences of L_1. We are now interested in this agent's inference from hearing the general term (*boat*) to the more frequent sub-kind (*motorboat*) as the referent. We obtain this variation by varying the prior over r_{COMMON} from 0.33 to 0.99. The y-axes show S_2's responses given the disjunctive observation $r_{\text{COMMON}} \vee r_{\text{UNCOMMON}}$. Dashes again correspond to spans in which violating HG with the disjunction *general or marked_specific* (e.g., *boat or canoe*) is this speaker's preferred utterance. Here again, α and the cost of disjunction influence this agent's choices. For instance, 1.06 is approximately the lowest α that generates this behavior, whereas high values of α make it more likely. Conversely, high disjunction costs again reduce the number of HG-violations by making them worthwhile for the speaker only if the listener's tendency to draw I-implicatures is very high.

The above simulations show that the model can capture not only implicature inferences but also the behavior that would normally be characterized as implicature blocking. This possibility arises from the recursive nature of the model, in which pragmatic speakers can anticipate the construals of pragmatic listeners and plan their own utterances accordingly. We should emphasize also that the Q-implicature and I-implicature scenarios represented by Figure 11 and Figure 12 are just particular examples. Since neither kind of implicature is a primitive of the model, but rather just emerges from other interactions, we are not limited to explanations that fall precisely into one of these two categories, which aligns well with our intuitions about the complex factors that guide HG violations in communication.

5.4 Parameter-Space Exploration

The above illustrative examples provide initial clues as to how our model characterizes Hurfordian and definitional uses from the speaker and listener perspectives. The goal of this section is to more thoroughly explore the space of options, with the goal of formulating a meta-theory of these behaviors and their relationships to the facts.

As we discussed at the start of this section, both uses require the speaker to be invested in some sense in communicating information about the lexicon. In a similar vein, the listener must have some lexical uncertainty, or at least be willing to defer to the speaker's preferences. Both of these forces are centered around the parameter β, which controls both speaker and listener behavior because of the recursive nature of the model. If β is set to 0, then the lexica are not distinguished from one another, and we end up with a system in which only information about the world is exchanged.

The above illustrative examples begin to show that β's relationship to the parameter α and the cost of disjunction are what steer the discourse participants to exclusivization or identification. Figure 13 gives a fuller picture of these dynamics using our larger lexicon setting (five atomic messages, four atomic states). The x-axis gives $\log(\beta/\alpha)$ for values of α and β between 0 and 14 inclusive, in increments of 1. The log-scale helps bring out the underlying relationships, and it also means that values above 0 are where β is bigger than α. The y-axis gives the cost of disjunction, ranging from 0 to 0.2, in increments of 0.01. The dots classify best inferences in this space of parameters, with green marking strictly dominant Hurford strategies, orange marking strictly dominant definitional strategies, and black marking cases where both strategies are strictly dominant (which is possible because we've reduced α and β to a single measure in order to visualize the space in two dimensions).

The picture that emerges from Figure 13 is relatively clear: definitional readings exist in a narrow space where $\beta > \alpha$ and disjunction costs are low. (Very small lexica and world spaces are somewhat more lenient about this, but it seems to hold in systems of sufficient complexity to allow meaningful adjustments to the messages.) As we said above, this makes intuitive sense: where costs are high, the disjunction has to be justified. Letting the two terms overlap reduces the justification, whereas exclusivizing provides justification with regard to α. However, the pressures of β also intrude. If β is high, then it might be worth paying the disjunction costs for the sake of teaching the listener about the lexicon, even if this reduces the amount of information conveyed per word.

In broad strokes, we can relate these ideas to Figure 13. At the bottom of the plot, where disjunction costs are low, we have, perhaps, a professor with ample time and not much to say; as we move right in this space, the professor becomes more invested in the "lingo", and hence definitional readings arise. At the top of the plot, the communications become time- or resource-constrained, and the pressures of HG become strong enough to compel exclusivization even with high β. From this perspective, it is clear why Hurfordian uses are more robust. They arise in a much wider range of contexts because they can easily survive high disjunction costs — exclusivization ensures genuine communicative value. In contrast,

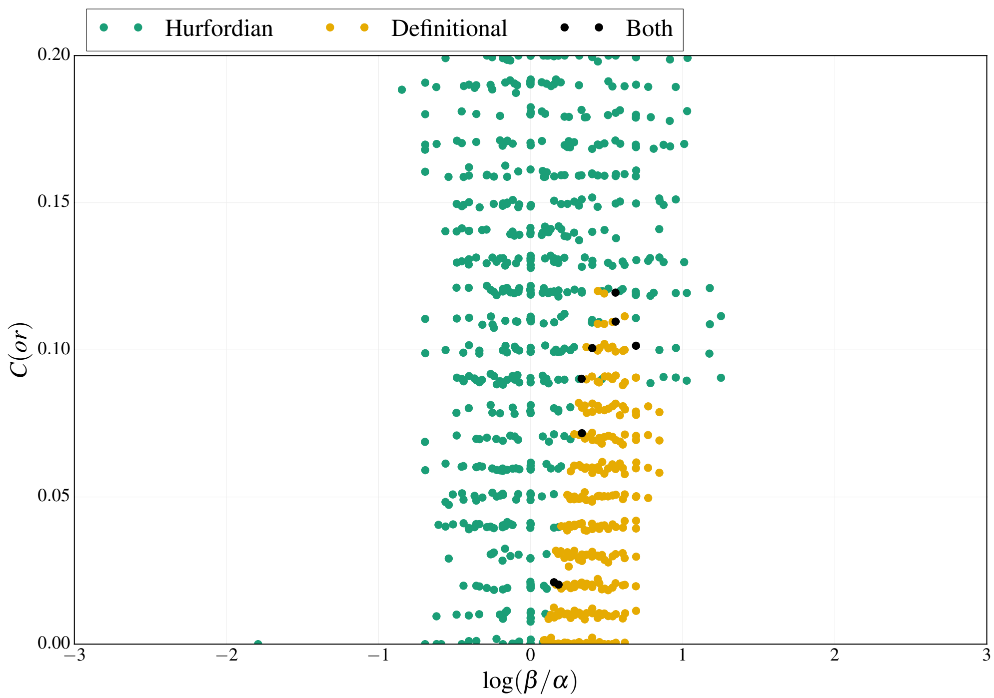


Figure 13: Hurfordian and definitional contexts with a large lexicon (five atomic lexical items, four atomic states).

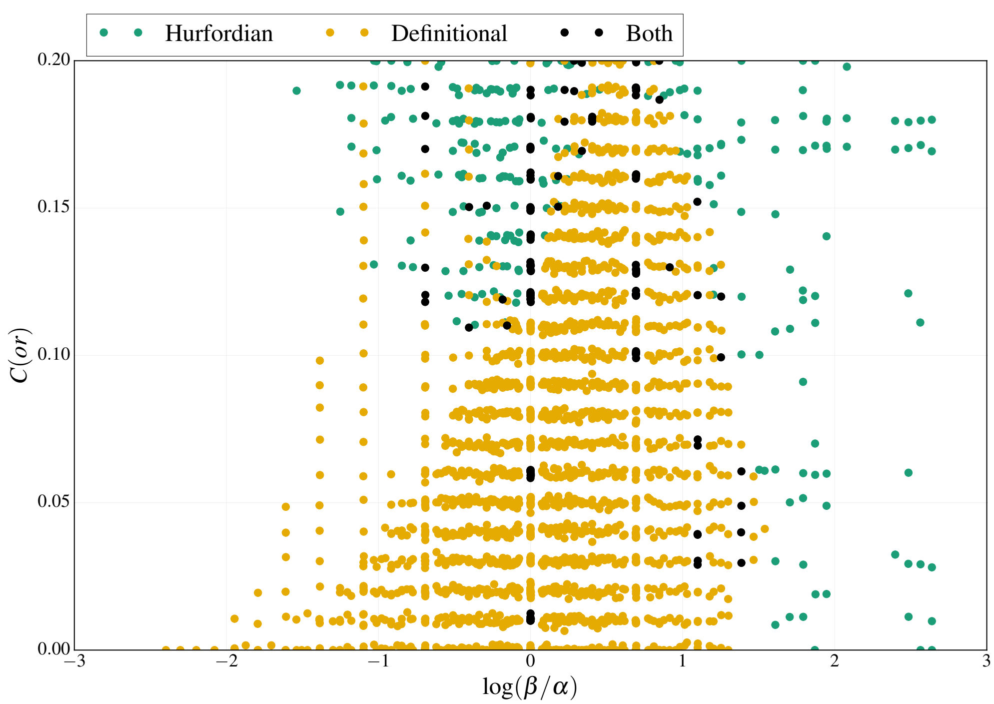


Figure 14: As in Figure 13, but with a focal point unknown word.

definitional readings exist mainly in the space of low disjunction and high β because the state information they convey is fully encoded in the first disjunct, making the full disjunction an inefficient way of conveying world information.

The quantitative picture also leads us to expect that there can be uncertainty about whether the listener should regard the disjunction as definitional or Hurfordian. The discourse participants might be in a blurry area in which small changes to the parameter settings push towards one inference or the other, with the difference between the best and next-best inferences relatively small. This can persist even when one of the words is unknown to the listener. (In such cases, the disjunctive meaning is just extremely general and will not do justice to the speaker's intentions.)

We saw in Section 5.1 that constraining the unknown word to have an atomic meaning — a meaning at the same level of specificity as the other lexical items — can greatly increase the strength of the definitional inference. Figure 14 shows that this is also reflected in the full parameter space. The figure is based on the same data as Figure 13 but with the focal point assumption constraining the space of lexica. The result is that definitional readings now exist in a much wider area of the parameter space: disjunction costs can be higher and β can be smaller than α. Thus, if a context supports this general constraint — for example, if the speaker is known to be trying to instruct the listener about words and concepts — then definitional readings should be more salient.

6 Conclusion

This paper synthesized and extended ideas from recent models of language production and construal, especially those of Smith et al. (2013) and Bergen et al. (2014), in order to provide a unified account of two seemingly conflicting inferences that disjunction supports — a pressure to exclusivize the disjuncts and a pressure to regard them as synonymous. From a single model of disjunctive semantics, coupled with general principles of pragmatic inference, a rich variety of uses of disjunction emerge: definitional interpretations (Section 5.1), Hurfordian exclusivization interpretations (Section 5.2), and defensive, implicature-blocking speaker behavior involving disjunction (Section 5.3). Each use within this rich variety traces to a specific set of relative priorities of the discourse participants with regard to communicating about the language itself, communicating about the world, and avoiding or tolerating undue message costs.

Our most fundamental contributions lie in the structure of the pragmatic model. We allow the speaker's communicative intentions to include linguistic preferences, and we allow listeners to make inferences about those preferences during the regular course of linguistic interactions. This structure is particularly important for exclusivization inferences, which do not affect truth conditions and so do not even emerge in most models. We hope this approach suggests new ways of detecting and understanding the secondary messages encoded in speakers' utterances, and that it can shed new light on the importance of communicating in language about language, and on related issues involving meta-linguistic negotiations and variable levels of perceived expertise about the language. More broadly, it could also serve as a way to connect pragmatic inference with phenomena relating to linguistic change and pedagogy, social identity, social hierarchies, and linguistic style.

7 References

Bergen, Leon; Noah D. Goodman; and Roger Levy. 2012. That's what she (could have) said: How alternative utterances affect language use. *Proceedings of the thirty-fourth annual conference of the Cognitive Science Society*, ed. by Naomi Miyake, David Peebles, and Richard P. Cooper, 120–125. Austin, TX: Cognitive Science Society.

Bergen, Leon; Roger Levy; and Goodman Noah D. 2014. Pragmatic reasoning through semantic inference. Ms., MIT, UCSD, and Stanford.

Brennan, Susan E., and Herbert H. Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22.1482–1493.

Camerer, Colin F.; Teck-Hua Ho; and Juin-Kuan Chong. 2004. A cognitive hierarchy model of games. *The Quarterly Journal of Economics* 119.861–898.

Chemla, Emmanuel. 2013. Apparent Hurford constraint obviations are based on scalar implicatures: An argument based on frequency counts. Ms., Ecole Normale Supérieure.

Chierchia, Gennaro; Danny Fox; and Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Maienborn et al., 2297–2332.

Clark, Herbert H. 1994. Managing problems in speaking. *Speech Communication* 15.243–250.

Clark, Herbert H. 1997. Dogmas of understanding. *Discourse Processes* 23.567–59.

Clark, Herbert H., and Meredyth A. Krych. 2004. Speaking while monitoring addressees for understanding. *Journal of Memory and Language* 50.62–81.

Clark, Herbert H., and Deanna Wilkes-Gibbs. 1986. Referring as a collaborative process. *Cognition* 22.1–39.

Degen, Judith, and Michael Franke. 2012. Optimal reasoning about referential expressions. *Proceedings of the 16th workshop on the semantics and pragmatics of dialogue*, ed. by Sarah Brown-Schmidt, Jonathan Ginzburg, and Staffan Larsson, Paris, 2–11.

Dillon, Brian; Charles Clifton; and Lyn Frazier. 2014. Pushed aside: Parentheticals, memory and processing. *Language, Cognition and Neuroscience* 29.483–498.

Fellbaum, Christiane (ed.) 1998. *WordNet: An electronic database.* Cambridge, MA: MIT Press.

Frank, Michael C., and Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science* 336.998.

Franke, Michael. 2009. *Signal to act: Game theory in pragmatics.* ILLC Dissertation Series. Institute for Logic, Language and Computation, University of Amsterdam.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition and logical form*. New York: Academic Press.

Giles, Howard; Nikolas Coupland; and Justine Coupland. 1991. Accommodation theory: Communication, context, and consequence. *Contexts of accommodation*, ed. by Howard Giles, Nikolas Coupland, and Justine Coupland, 1–68. Cambridge: Cambridge University Press.

Goodman, Noah D., and Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science* 5.173–184.

Hearst, Marti A. 1992. Automatic acquisition of hyponyms from large text corpora. *Proceedings of the 15th international conference on computational linguistics*, 539–545. Nantes: Association for Computational Linguistics.

Hirschberg, Julia. 1985. *A theory of scalar implicature*. University of Pennsylvania dissertation.

Horn, Laurence R. 1989. *A natural history of negation*. Chicago: University of Chicago Press.

Hurford, James R. 1974. Exclusive or inclusive disjunction. *Foundations of Language* 11.409–411.

Jäger, Gerhard. 2007. Game dynamics connects semantics and pragmatics. *Game theory and linguistic meaning*, ed. by Ahti-Veikko Pietarinen, 89–102. Amsterdam: Elsevier.

Jäger, Gerhard. 2012. Game theory in semantics and pragmatics. In Maienborn et al., 2487–2425.

Kao, Justine T.; Leon Bergen; and Noah D. Goodman. 2014a. Formalizing the pragmatics of metaphor understanding. *Proceedings of the 36th annual meeting of the Cognitive Science Society*, 719–724. Wheat Ridge, CO: Cognitive Science Society.

Kao, Justine T.; Jean Y. Wu; Leon Bergen; and Noah D. Goodman. 2014b. Non-literal understanding of number words. *Proceedings of the National Academy of Sciences* 111.12002–12007.

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

Levy, Roger; Leon Bergen; and Goodman Noah D. To appear. *Roses and flowers*: An informativeness implicature in probabilistic pragmatics. *Proceedings of the 24th conference on semantics and linguistic theory*. Ithaca, NY: CLC Publications.

Levy, Roger, and Carl Pollard. 2001. Coordination and neutralization in HPSG. *Proceedings of the 8th international conference on Head-Driven Phrase Structure Grammar*, ed. by Frank Van Eynde, Lars Hellan, and Dorothee Beermann, 221–234. Stanford, CA: CSLI.

LEWIS, DAVID. 1969. *Convention.* Cambridge, MA: Harvard University Press.

LIBERMAN, MARK. 2005. Questioning reality. http://itre.cis.upenn.edu/~myl/languagelog/archives/001837.html.

MAIENBORN, CLAUDIA; KLAUS VON HEUSINGER; and PAUL PORTNER (eds.) 2012. *Semantics: An international handbook of natural language meaning*, vol. 3. Berlin: Mouton de Gruyter.

MICHEL, JEAN-BAPTISTE; YUAN KUI SHEN; AVIVA PRESSER AIDEN; ADRIAN VERES; MATTHEW K. GRAY; THE GOOGLE BOOKS TEAM; JOSEPH P. PICKETT; DALE HOIBERG; DAN CLANCY; PETER NORVIG; JON ORWANT; STEVEN PINKER; MARTIN A. NOWAK; and EREZ LIEBERMAN AIDEN. 2011. Quantitative analysis of culture using millions of digitized books. *Science* 331.176–182.

POTTS, CHRISTOPHER. 2005. *The logic of conventional implicatures.* Oxford: Oxford University Press.

POTTS, CHRISTOPHER. 2012. Conventional implicature and expressive content. In Maienborn et al., 2516–2536.

RUSSELL, BENJAMIN. 2012. *Probabilistic reasoning and the computation of scalar implicatures.* Providence, RI: Brown University dissertation.

SCHELLING, THOMAS C. 1960. *The strategy of conflict.* Cambridge, MA: Harvard University Press.

SIMONS, MANDY. 2001. Disjunction and alternativeness. *Linguistics and Philosophy* 24.597–619.

SINGH, RAJ. 2008. On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy* 31.245–260.

SMITH, NATHANIEL J.; NOAH D. GOODMAN; and MICHAEL C. FRANK. 2013. Learning and using language via recursive pragmatic reasoning about other agents. *Advances in neural information processing systems 26*, ed. by Christopher J. C. Burges, Leon Bottou, Max Welling, Zoubin Ghahramani, and Kilian Q. Weinberger, 3039–3047. Lake Tahoe, CA: Neural Information Processing Systems Foundation.

SNOW, RION; DANIEL JURAFSKY; and ANDREW Y. NG. 2005. Learning syntactic patterns for automatic hypernym discovery. *Advances in neural information processing systems 17*, ed. by Lawrence K. Saul, Yair Weiss, and Léon Bottou, 1297–1304. Cambridge, MA: MIT Press.

STILLER, ALEX; NOAH D. GOODMAN; and MICHAEL C. FRANK. 2011. Ad-hoc scalar implicature in adults and children. *Proceedings of the 33rd annual meeting of the Cognitive Science Society*, ed. by Laura Carlson, Christoph Hoelscher, and Thomas F. Shipley, 2134–2139. Austin, TX: Cognitive Science Society.

Syrett, Kristen, and Todor Koev. 2014. Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics*.

van Tiel, Bob; Emiel van Miltenburg; Natalia Zevakhina; and Bart Geurts. 2014. Scalar diversity. *Journal of Semantics*.

Tonhauser, Judith; David I. Beaver; Craige Roberts; and Mandy Simons. 2013. Towards a taxonomy of projective content. *Language* 89.66–109.

Vogel, Adam; Max Bodoia; Christopher Potts; and Dan Jurafsky. 2013. Emergence of Gricean maxims from multi-agent decision theory. *Human language technologies: The 2013 annual conference of the North American chapter of the Association for Computational Linguistics*, 1072–1081. Stroudsburg, PA: Association for Computational Linguistics.