Resolving Quantity- and Informativeness-Implicature in Indefinite Reference

Till Poppels and Roger Levy *

Department of Linguistics
University of California, San Diego
{tpoppels,rlevy}@ucsd.edu

Abstract

A central challenge for all theories of conversational implicature (Grice, 1957, 1975) is characterizing the fundamental tension between Quantity (Q) implicature, in which utterance meaning is refined through exclusion of the meanings of alternative utterances, and Informativeness (I) implicature, in which utterance meaning is refined by strengthening to the prototypical case (Atlas & Levinson, 1981; Levinson, 2000). Here we report a large-scale experimental investigation of Q-I resolution in cases of semantically underspecified indefinite reference. We found strong support for five predictions, strengthening the case for recent rational speaker models of conversational implicature (Frank & Goodman, 2012; Degen, Franke, & Jäger, 2013): interpretational preferences were affected by (i) subjective prior probabilities (Informativeness), (ii) the polarity and (iii) the magnitude of utterance cost differentials (Quantity), (iv) the felicity conditions of indefinite NPs in English, and (v) the 'relatability' of X and Y.

1 Introduction

1.1 The phenomenon

In transitive sentences of the form 'The X V-ed a Y' the relationship between X and Y remains semantically unspecified and must be inferred (Horn, 1984):

(1) a. The man injured a child. +>[1] not his own child
b. The man broke a finger. +> his own finger

In these examples the indefinite reference to Y is ambiguous between a relational reading, where Y is X's own (as in 1b; we will be referring to this as the 'OWN interpretation'), and a non-relational reading, where Y does not belong to X (as in 1a; henceforth referred to as 'OTHER'S'). In the way this OWN/OTHER'S ambiguity is resolved, (1) illustrates the tension between two major constraints on language production that have been recognized as driving extrasemantic inferences since Grice (1957): Quantity and Informativeness. In Grice's original proposal, these two constraints were captured in the Maxim of Quantity, which instructs speakers to be brief, while also being informative. Subsequent efforts to reduce redundancy in the Gricean

*This work has benefited from feedback from David Barner, Andrew Kehler, Dan Lassiter, members of the Computational Psycholinguistics Lab at UCSD, and audiences at XPRAG 2015 and AMLaP 2015. We gratefully acknowledge support from a UCSD Dean's Fellowship to TP and from a UCSD Academic Senate research grant, NSF grant BCS-1456081, and an Alfred P. Sloan Research Fellowship to RL.

[1] '+>' is short for 'implicates'.

Maxims, most notably by Levinson (2000) and Horn (1984), have singled out Quantity and Informativenss as key sources of Gricean inferences.[2]

The effect of Quantity can be witnessed in 1a, which is completely analogous to standard cases of Quantity (Q-) implicature: The OTHER'S interpretation of *The man injured **a** child* is Q-implicated through exclusion of the meaning of a similarly brief, yet unambiguous and thus more informative, alternative utterance for conveying OWN: *The man injured **his** child.* By the same reasoning we would expect an analogous Q-implicature from 1b, yet the opposite inference appears to arise. This has been argued to reflect the effect of Informativeness (I-) implicature (Atlas & Levinson, 1981), i.e. the strengthing of utterance meaning to the "stereotypical meaning, use, or situation" (Horn, 2004, p. 16).

Thus, the puzzle in 1 illustrates the fact that the tension between Quantity and Informativeness is resolved differently across different utterances. The present paper addresses *how* and *when* this Q/I tension is resolved in favor of Q (as in 1a) or I (as in 1b).

1.2 The Rational Speech Act model

At the center of the Gricean program is the Cooperative Principle, which licenses the assumption (among others) that the speaker's utterance of choice aims to strike an optimal balance between being maximally informative (I) and efficient (Q) compared to its alternatives. This core idea has been formalized in recent Bayesian (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013) and game-theoretic (Degen et al., 2013) models, in which interlocutors maintain probabilistic beliefs about each other's knowledge and communicative goals, and rely on these beliefs to reason iteratively about each other's choices. Here, we focus on the Rational Speech-Act (RSA) model, which models the interpretational preferences of a pragmatic listener L_1, whose speaker model (S_1) is based on a hypothetical literal listener (L_0) who interprets utterances (u) according to their literal semantics and a prior distribution over meanings (m):[3]

$$P_{L1}(m|u) \propto P_{S1}(u|m)P(m) \tag{i}$$

$$P_{S1}(u|m) \propto \exp(\lambda[\log(P_{L_0}(m|u)) - D(u)]) \tag{ii}$$

$$P_{L_0}(m|u) \propto \mathcal{L}(u,m)P(m) \tag{iii}$$

This iterative reasoning process grounds out in the mutually known 'lexicon' $\mathcal{L}$, which maps utterances to meanings that are consistent with it:

$$\mathcal{L}(u,m) = \begin{cases} 1 & \text{if } m \in [\![u]\!] \\ 0 & \text{otherwise} \end{cases} \tag{iv}$$

While L_0 and L_1 reason about potential meanings of a given utterance, S_1 models the tension between being optimally brief and informative: this 'Gricean speaker' chooses (softmax with

[2]Note that Horn refers to the Informativeness constraint as R, and that Levinson's taxonomy includes an additional component, M, which roughly maps onto Grice's Maxim of Manner. For present purposes, we restrict our discussion to Q and I, however, since it is the resolution of these two interpretational forces that we are looking to explain.

[3]This formalization goes back to Frank and Goodman (2012) and Goodman and Stuhlmüller (2013), although in the original proposal the literal listener does not use the prior distribution, so that $P_{L0}(m|u) \propto \mathcal{L}(u,m)$. In line with subsequent extensions of the model (Degen, Tessler, & Goodman, 2015; Lassiter & Goodman, 2015), we do attribute knowledge of the prior to L_0, in order to enable S_1 to capitalize on shared world knowledge when choosing the optimal utterance for conveying her meaning.

hardness parameter λ) between alternative utterances for expressing her intended meaning by weighing the cost $D(u)$ of each utterance u against the surprisal that her intended meaning m would have in the posterior distribution $P_{L_0}(m|u)$ of her listener model L_0 (Eq. ii). This is where alternative utterances exert what we will call their 'scalar pressure': since speakers generally choose utterances that are maximally informative about their intended meaning, every utterance that is *not* chosen by the speaker exerts some amount of interpretational pressure away from the meaning it encodes. Crucially, however, S_1 is also constrained by costs, which we assume are monotonically decreasing in utterance brevity. Thus, L_1 can explain away the speaker's choice not to use a particular alternative utterance if that utterance is more prolix than her utterance of choice. This means that when an utterance u is compatible with a meaning m for which a more precise utterance u' is also available, the less costly u' is the more scalar pressure it exerts against u being interpreted as m. This notion of 'scalar pressure' is central to the present study and we will appeal to it repeatedly in the remainder of the paper.

In order to apply the RSA model to the utterances such as 1, we assume a minimal lexicon that contains the (ambiguous) utterance in question along with its unambiguous alternatives—schematically represented as $\{a$; *self's*; *someone else's*$\}$[4]—and a set of two possible meanings: OWN and OTHER'S. We further assume a cost function, which assigns production costs to utterances based on the number of syllables they comprise: $D(a) = D(\textit{self's}) = 1$ and $D(\textit{someone else's}) = 4$.[5] Table 1 summarizes these assumptions. Figure 1 graphically depicts the lexical meaning relationship among alternative utterances; the fainter depiction of *someone else's* reflects its higher cost, suggesting that it should exert less scalar pressure on a toward an OWN interpretation than *self's* does toward an OTHER'S interpretation. Throughout the present paper, we further assume that $\lambda = 1$, recovering a Luce choice rule (Luce, 1959), which has used to model human decisions across a diverse set of domains (Sutton & Barto, 1998). Greater values for λ produce more polarized production preferences, although they are qualitatively robust across different parameter settings.

utterance (u)	$\mathcal{L}(u$, OWN)	$\mathcal{L}(u$, OTHER'S)	$D(u)$
a	1	1	1
self's	1	0	1
someone else's	0	1	4

Table 1: Assumed 'lexicon' and utterance costs

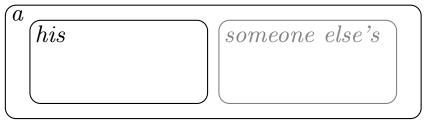

Figure 1: Graphical depiction

With these assumptions in place, Figure 2 shows the model's posteriors at successive stages of the iterative reasoning process as a function of the prior distribution over meanings $P(m)$. Consider first the literal listener (left panel). Owing to the 'lexicon' just described, L_0's probability of OTHER'S is 0 given *self's* and 1 given *someone else's*, whereas the interpretation of the ambiguous utterance (solid line graph) is driven entirely by the prior, since it is (by definition) compatible with either meaning. Consequently, *self's* cannot be used to convey OTHER'S to L_0 (therefore, $P_{S1}(self's|\text{OTHER'S}) = 0$), leaving a to compete with *someone else's*, since both of these alternatives are literally compatible with S_1's intended meaning (top panel). In choosing between these two utterances, S_1 is drawn towards *someone else's* because it is guaranteed to be interpreted correctly (by L_0), and therefore maximally informative. However, the relative

[4]Throughout the paper we use the umbrella term '*self's*' to refer to the forms that convey OWN unambiguously, which include *his*, *her*, *its*, *our*, *your*, *their*, and *my*, all of which are monosyllabic.

[5]The monotonicity, not the exact numbers, of this cost profile is crucial for the qualitative predictions we address here. In principle, the psychological cost of an utterance is likely not reducible to its length, and exactly what determines production costs remains an open question. We will not address this issue here, although we do propose an extension to this cost function in Section 1.3.

brevity of a is attractive as well, and its informativeness increases with the prior probability of OTHER'S. Thus, S_1 will bother to use the unambiguous utterance to convey OTHER'S only when the prior probability of that meaning is below a certain threshold (in the present example the cross-over point is at $P(\text{OTHER'S}) = 0.05$. Crucially, this contrasts with the decision facing a speaker who wants to convey OWN, since in that case both literally compatible utterances are equally inexpensive ($D(a) = D(his) = 1$). In that scenario (bottom panel), the speaker will be drawn to the unambiguous utterance unless, of course, the $P(\text{OTHER'S}) = 0$, in which case the prior rules out OTHER'S, and a and *self's* are equally informative, so that $P_{S1}(a|\text{OWN}) = P_{S1}(\textit{self's}|\text{OWN}) = .5$.

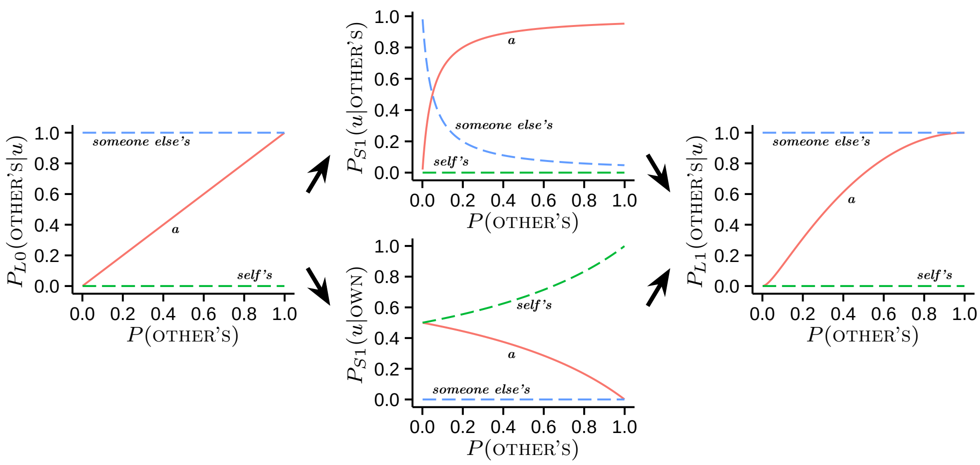

Figure 2: Posterior distribution of the literal listener (left), the 'Gricean speaker' (center), and the pragmatic listener (right). The speaker's intention is to convey OTHER'S in the top panel, and OWN in the bottom one. The solid line represents the ambiguous utterance.

The pragmatic listener (right panel) then attempts to infer S_1's intention by reverse-engineering the production process, and integrating S_1's posterior with prior probabilities. As a result, her interpretation of the ambiguous utterance is skewed towards OTHER'S relative to the prior. However, since the scalar pressure driving this Q-implicature is proportional to the Informativeness benefit of using *self's* instead of a, the effect of Q can be overturned by I if the prior tends strongly towards OWN.

Note the RSA model closely captures traditional definitions of Quantity implicature as alternative-based inferences and Informativeness implicature as the strengthening of utterances to the stereotypical meaning. Crucially, however, it goes beyond these definitions by providing the necessary machinery for explaining the trade-off between these two interpretational forces in a principled way, by casting it as the prior-likelihood trade-off in Bayesian inference. As a result, the model is capable of generating precise predictions about Q/I resolution in sentences like 1, which are desribed in the next section.

1.3 Predictions

We conducted a large-scale forced-choice experiment to investigate interpretational preferences in sentences like (1). Using mixed logit regression, we test five predictions about the way the Q/I tension is resolved in these sentences. Three of these predictions follow directly from the RSA model and are illustrated in Figure 3. Two additional predictions are motivated independently and explore potential limitations of the RSA model in accounting for Informativeness-driven inferences.test

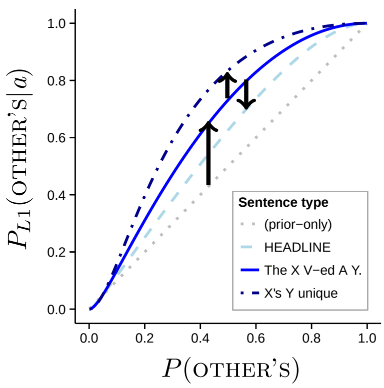

Figure 3: Predictions

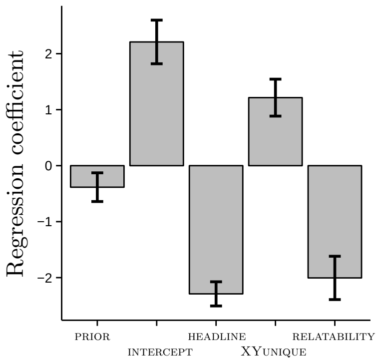

Figure 4: Regression coefficients and Standard Errors

Predictions 1 and 2: Event priors and baseline *Q*-implicature. Since rational-speaker models generally assume that speakers will capitalize on their common ground with their addressee, mutual knowledge about the world represents the starting point for pragmatic reasoning according to these models. Within the RSA model, interpretations are partly determined by, and should therefore track, prior probabilities, which corresponds to the general monotonicity of the prior–posterior relationship in all lines of Figure 3. But relative to the prior, interpretations should be skewed towards OTHER'S due to the scalar pressure from *self's*: in Figure 3, all model predictions are above the $x = y$ line.

Prediction 3: Reduced effect of *Q* in headline cases. Manipulating the utterance cost profile should modulate the effect of Q. In particular, if we reduce the cost of the ambiguous utterance then it should reduce overall scalar pressure and bring the L_1 posterior closer to the prior. Consider the following contrast:

(2) a. The man injured a child.
b. Man injured child.

The HEADLINE-like style of 2b affords an ambiguous utterance with no explicit determiner—we denote this option with $\emptyset$. If utterance costs are monotonic in length, then $\emptyset$ is cheaper than either unambiguous alternative: $D(\emptyset) < D(\textit{self's}) < D(\textit{someone else's})$. As illustrated in Figure 3, RSA predicts for this modified cost profile an overall weaker effect of Q-implicature in headline-like sentences. Since the original effect of Q was a skew toward OTHER'S interpretations, headline-like sentences should be OWN-skewed relative to their full-sentence counterparts.

Prediction 4: Non-uniqueness in indefinite reference. Hawkins (1991) argues that the use of the indefinite determiner in English is infelicitous whenever its reference is uniquely satisfied within the domain:

(3) a. # a brightest student
b. # a president of the United States

This generalization makes an interesting prediction about the sentences we are considering ('The X V-ed a Y'): in cases where the semantics of X and Y are such that a typical X has only one Y, the non-uniqueness requirement of the indefinite is violated on an OWN interpretation, rendering the indefinite reference infelicitous given that OWN is intended. If this felicity condition is mutually known and constrains the speaker's utterance choice, we would expect 4b to be less likely to receive an OWN interpretation than 4a:

(4) a. The man broke a finger.
b. The man broke a nose.

This prediction can be accommodated within the RSA model under the auxiliary assumption that D be made a 2-place function of u, m pairs (Jäger, 2012), such that not only utterance brevity but also 'felicity' of specific form–meaning pairs contribue to costs. Let the value of the new $D(u, m)$ be the value of $D(u)$ from Table 1 in all cases except for $u = a$ and $m =$ OWN in contexts where X and Y are such that a typical X possesses only a single Y (as in 4b), in which case $D(a, own) > D(self's, own)$. In these contexts, S_1 will be disincentivized from using the ambiguous utterance when aiming for an OWN interpretation, and correspondingly shifts some probability mass from OWN to OTHER'S in L_1's posterior.

Prediction 5: The real-world 'relatability' of event participants. The materials used in the present study include event descriptions that vary with respect to the 'relatability' of the event participants, exemplified by the following contrast, where *father* and *son* are intuitively more relatable than *man* and *child*:

(5) a. The man injured a child.
b. The father injured a son.

Precisely characterizing relatability is beyond the scope of this paper; we informally take it as an index of the degree to which the description of one NP brings into mind another referent with which the other NP could be identified. Relatability seems closely related to bridging (Clark, 1975), as in the following example, due to Kehler (in prep):

(6) I almost bought a car today, but...
a. ... *the engine* was too noisy.
b. # ... *the TV* was blurry.
c. ## ... *the stapler* was broken.

The relationship between the car and the entity denoted by the italicized definite NP is semantically unspecified and must be inferred. That inference appears more natural, however, when the entity in question is highly relatable to the car (cf. *engine* in 6a) than when the two are less relatable (as in 6b and 6c). By analogy to such bridging inferences, one may expect that sentences with highly relatable NPs, such as (5b), tend more strongly towards an OWN interpretation, since on that reading the two entities *are* related. It is important to note, however, that the RSA model has no mechanism for predicting such effects at present. We will return to this issue in the Discussion.

2 Methods

Participants. 1885 native speakers of English were recruited via Amazon.com's Mechanical Turk to complete a single-trial forced-choice experiment. A UniqueTurker script[6] ensured that each participant could complete the survey only once.

Materials. 53 sentences of the form *The X V-ed a Y* were designed to vary widely with respect to prior OWN/OTHER'S probability, the relatability of X and Y, as well as the number of Ys a typical X possesses. Each of those X-V-Y sentences was matched with a corresponding newspaper headline-like version of the form *X V-ed Y*.

Procedure. Participants were presented with a single-trial forced-choice task, which was implemented using the online survey platform Qualtrics. The answer choices identified OWN and OTHER'S readings of the sentences and were presented in a random order.

Prior norming. The prior OWN/OTHER'S probability of each of the 53 items was estimated in a separate norming experiment. We recruited 885 participants from Mechanical Turk who were presented with five questions of the form 'How likely is an X to V his own Y compared to V-ing someone else's Y?', each corresponding to one of the 53 events described in the main experiment. Using a slider, participants chose between '100% likely to V his/her/its own Y' and '100% likely to V someone else's Y', describing OWN and OTHER'S events, respectively.

3 Results

The interpretational preferences measured in the main experiment ranged from 100% OWN to 100% OTHER'S, and estimated prior probabilities from 80.49% OWN to 88.59% OTHER'S. A logistic regression analysis with mixed effects tested whether interpretations varied as a function of prior probabilities, whether the sentence was presented as a full sentence or a headline-like version (HEADLINE), whether or not a typical X had only one Y (XYUNIQUE), and the RELATABILITY of X and Y. Values for XYUNIQUE and RELATABILITY were hand-coded by the first author. Prior estimates p from the norming experiment were scaled to range from 0 to 1, centered around 0.5, and logit-transformed (PRIOR = logit $\left[\frac{0.8\times p+10}{100}\right]$). All other factors were treatment-coded categorical predictors and HEADLINE was added as a random by-item effect.[7] The results are overall consistent with our predictions (Fig. 4): The prior had a numerical effect in the predicted direction, although this trend did not reach significance ($\beta = -.38, p = .13$). The significant INTERCEPT indicates the predicted OTHER'S-skew ($\beta = 2.21, p < .001$), which was enhanced where X's Y was unique ($\beta = 1.22, p < .001$). HEADLINEs were significantly more likely to receive OWN interpretations than their full-sentence counterparts ($\beta = -2.29, p < .001$), and so were sentences with highly relatable nouns compared to those with less relatable nouns ($\beta = -2.01, p < .001$).

[6] uniqueturker.myleott.com

[7] The full formula was `response ~ prior + XYunique + relatability + headline + (1 + headline|item)`.

4 Discussion

Support for the RSA model. The overall OTHER'S-skew (as indicated by the significant negative intercept) reflects the predicted baseline effect of Q-implicature: because *self's* is lower-cost than *someone else's*, it exerts more scalar pressure on the interpretation of the ambiguous *a*, pushing the latter towards an OTHER'S interpretation. Furthermore, the within-item HEADLINE manipulation confirms RSA's prediction that by lowering the cost of the ambiguous utterance, scalar pressure on it and thus the OTHER'S-skewing Q effect would be ameliorated. Together, these results provide strong support for the RSA model.

Event priors. Since we expected interpretations to track prior probabilities, it is worth speculating why the effect of the prior did not reach significance. Notice that on a strict interpretation of the RSA model as Bayesian inference, with the posterior distribution defined over communicative intentions, the corresponding priors should be defined over intentions as well. Our norming experiment, on the other hand, measured people's expectations about events, not event descriptions. While the probability of an event may be correlated with its probability of being mentioned, that correlation is likely to be noisy: extremely unlikely events may be more remarkable and therefore more likely to be talked about than highly predictable events. If this reasoning is correct, our prior estimates may not have the correct "currency," which may explain why they did not affect interpretations significantly. This contrast between event priors and 'intention priors' is mirrored in the definition traditionally given to I-implicature, as the strengthening of utterance meaning to the "stereotypical *use* (...) or *situation*" (Horn, 2004, p. 16; emphasis added).

Encoding felicity conditions as form-meaning costs. The prediction that the interpretation of *The X V-ed a Y* would tend towards OTHER'S in cases where X has only one Y is supported by our data. This prediction was based on Hawkins (1991)'s observation that indefinite reference in English tends to be infelicitous if it is known to be uniquely satisfied. To derive this prediction from within the RSA architecture, we encoded the non-uniqueness requirement as a soft constraint on production preferences through an additional cost function that assigned an extra 'felicity cost' whenever the speaker's intention was to elicit an OWN interpretation and X's Y was unique. Encoding felicity conditions as form-meaning costs captures the intuition that these constraints are part of interlocutors' mutual knowledge about the language they use to communicate.

The effect of relatability: *Q* or *I*? One may be tempted to explain the relatability effect by invoking *ad-hoc* scalar (Q-) implicature (Hirschberg, 1985) about referring expressions. On that view, highly relatable entities drive interpretations towards OWN because the speaker's use of referring expressions that are more specific than some of its scale mates (e.g. *father* and *son* as opposed to *man* and *child*) triggers a search for the purpose of that choice. Crucially, however, this explanation requires the auxiliary assumption that more specific referring expressions are in general more costly than their less specific competitors, which would have to be motivated independently in order to avoid circularity.[8]

Instead, the prediction that relatability would drive interpretations towards OWN was motivated by analogy to bridging inferences (Clark, 1975), in which reference resolution is biased towards entities that are associated with previously mentioned entities (recall that 'the engine'

[8] In principle, such production costs could correspond to 'lexical retrieval effort' if, for example, basic-level category labels, such as *man*, are easier to retrieve from memory than alternatives like *father*.

in 6a is strengthened to 'the engine of that car'). According to Prince and Cole (1981), such bridging inferences arise when entities associated with a previously mentioned entity are more 'inferrable' than their competitors, such as the engine of the previously mentioned car compared to other engines in the world. If this analogy holds, the relatability effect in our data reflects an Informativeness implicature that is based on the interlocutors' mutual real-world knowledge about the relatability of event participants.

This raises an interesting point about the nature of Informativeness-driven inferences: since I-implicatures are based on shared real-world knowledge, their *content* may be determined by entirely non-intentionalist inference mechanisms for making sense of the world, rather than iterative reasoning between interlocutors. Consider in this context the following example, due to Cohen and Kehler (submitted):

(7) The manager fired the employee who came in late every day last week.

Cohen and Kehler argue that the 'causal attribution' inference that this utterance invites—that the employee was fired *because* of being tardy—would arise in much the same way if the situation was perceived directly, rather than described linguistically: if you saw an employee who you know to be tardy getting fired, you may come to the same conclusion, although in that case there is no communicative intent to be inferred. Crucially, however, if the relevant information *is* conveyed intentionally (e.g. by uttering 7), the listener may reason that the speaker intended to elicit that inference, which Cohen and Kehler therefore call a 'conversational eliciture'. In that case, elicitures show a 'focal point effect' (Schelling, 1980), by which the inference is strengthened iteratively, although its content is fixed by ego-centric, non-intentionalist inference mechanisms. With respect to our X-V-Y, that would mean that the listener is drawn towards OWN when X and Y are highly relatable, and that tendency is strengthened by the assumption that the speaker must have anticipated that inference and wanted it to be drawn. Levinson appears to consider the possibility that Informativeness-based inferences may have non-intentionalist components when he likens them to 'inferences to the best explanation' in scientific contexts (cf. Thagard, 2000, 2007). Future work will be needed to further explore the role that semantic relatability, specifically, and non-intentionalist inferences, in general, play in language comprehension. As a first step towards that goal, future studies should aim to find reliable and fine-grained ways of quantifying relatability, such as corpus-based co-occurrence measures.

5 Conclusion

We set out to investigate the resolution of Quantity and Informativeness implicature in semantically underspecified cases of indefinite reference. We have demonstrated that the resolution of these antinomic interpretational forces involves several interacting factors, many of which follow straightforwardly from rational-speaker accounts of interpretation. Knowledge from various domains, such as semantic, pragmatic, social, and real-world knowledge, plays a role in the inferential mechanism that drives the interpretation of utterances. We have argued that developing predictive theories that explain exactly how interpretational inferences are drawn is an important theoretical goal and suggested that a complete account may ultimately require a non-intentionalist component. Finding a place for such a component (if it exists) within rational-speaker accounts may extend their explanatory reach and further our understanding of the cognitive machinery behind language comprehension.

References

Atlas, J. D. & Levinson, S. C. (1981). *It*-clefts, informativeness, and logical form: radical pragmatics (revised standard version). *Radical pragmatics*, *1*, 1–61.

Clark, H. (1975). Bridging. In *Proceedings of the 1975 workshop on theoretical issues in natural language processing*. Stroudsburg, PA, USA: Association for Computational Linguistics.

Cohen, J. & Kehler, A. (submitted). Conversational Elicitures.

Degen, J., Franke, M., & Jäger, G. (2013). Cost-based pragmatic inference about referential expressions. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the annual meeting of the cognitive science society* (pp. 376–381).

Degen, J., Tessler, M. H., & Goodman, N. D. (2015). Wonky worlds: Listeners revise world knowledge when utterances are odd. In *Preceedings of the 37th annual conference of the cognitive science society*.

Frank, M. C. & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998.

Goodman, N. D. & Stuhlmüller, A. (2013). Knowledge and implicature: modeling language understanding as social cognition. *5*(1), 173–184.

Grice, H. P. (1957). Meaning. *The Philosophical Review*, *66*(3), 377–388.

Grice, H. P. (1975). Logic and conversation. In P. Cole & G. Morgan (Eds.), *Syntax and semantics III* (pp. 41–58). London: Academic Press.

Hawkins, J. A. (1991). On (in)definite articles: implicatures and (un) grammaticality prediction. *Journal of Linguistics*, *27*(2), 405–442.

Hirschberg, J. (1985). A Theory of Scalar Implicature (Natural Languages, Pragmatics, Inference). *Dissertations available from ProQuest*. Retrieved from http://repository.upenn.edu/dissertations/AAI8603648

Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context: linguistic applications* (pp. 11–42). Washington: Georgetown University Press.

Horn, L. R. (2004). Implicature. In L. R. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 3–28). Blackwell Publishing Ltd.

Jäger, G. (2012). Game theory in semantics and pragmatics. In C. Maienborn, P. Portner, & K. von Heusinger (Eds.), *Sematics: an international handbook of natural language meaning* (Vol. 3, pp. 2487–2516). Berlin: de Gruyter.

Kehler, A. (in prep). Reference in Discourse. In S. Lappin & C. Fox (Eds.), *Handbook of contemporary semantics - second edition* (pp. 851–890). Wiley-Blackwell.

Lassiter, D. & Goodman, N. D. (2015). Adjectival vagueness in a Bayesian model of interpretation. *Synthese*. Retrieved from http://web.stanford.edu/%7B~%7Ddanlass/Lassiter-Goodman-adjectival-vagueness-draft.pdf

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

Luce, R. D. (1959). Individual Choice Behavior a Theoretical Analysis.

Prince, E. F. & Cole, P. (1981). Toward a taxonomy of given-new information. (pp. 223–255).

Schelling, T. C. (1980). *The Strategy of Conflict*. Harvard University Press.

Sutton, R. S. & Barto, A. G. (1998). *Reinforcement learning*. MIT Press.

Thagard, P. (2000, November). *Coherence in thought and action*. MIT Press.

Thagard, P. (2007). Coherence, truth, and the development of scientific knowledge. *Philosophy of Science*, *74*(1), 28–47.