A rational speech-act model of projective content

Ciyang Qing, **Noah D. Goodman**, **Daniel Lassiter**
{qciyang, ngoodman, danlassiter}@stanford.edu
Stanford University

Abstract

Certain content of a linguistic construction can project when the construction is embedded in entailment-canceling environments. For example, the conclusion that John smoked in the past from the utterance *John stopped smoking* still holds for *John didn't stop smoking*, in which the original utterance is embedded under negation. There are two main approaches to account for projection phenomena. The semantic approach adds restrictions of the common ground to the conventional meaning. The pragmatic approach tries to derive projection from general conversational principles. In this paper we build a probabilistic model of language understanding in which the listener jointly infers the world state and what common ground the speaker has assumed. We take change-of-state verbs as an example and model its projective content under negation. Under certain assumptions, the model predicts the projective behavior and its interaction with the *question under discussion* (QUD), without any special semantic treatment of projective content.

Keywords: Presupposition; projection; Bayesian pragmatics

Introduction

> "How am I to get in?" asked Alice again, in a louder tone.
> "*Are* you to get in at all?" said the Footman. "That's the first question, you know." — Carroll (1866), ch. VI

Courtroom drama, political misinformation, and ordinary misunderstandings often revolve around *presupposition*, a backgrounded aspect of meaning with a complex logic and communicative function. Famously, presuppositions can be used manipulatively, as in the classic loaded question "Have you stopped beating your wife?", or the sly reporter's "When did you become aware that this policy was a failure?". At the same time, presuppositions serve to streamline conversation by allowing interlocutors to convey multiple pieces of information simultaneously. Alice's question "How am I to get in?" efficiently indicates both Alice's assumption that she will get in, and her wish to know how to enter.

In addition to their important communicative role, presuppositions are interesting because they seem to flout some of the most basic rules of logic. For example, from *John danced* we can infer *John moved*, but we cannot infer this from *John didn't dance*—the inference that *John moved* is canceled by negation. In contrast, from both *John stopped smoking* and *John didn't stop smoking* we are likely to infer that *John used to smoke*—this inference is said to *project* over negation. In natural language semantics and pragmatics, an inference that survives an operator that usually cancels inferences is called *projective content* of the sentence under that operator. Change-of-state verbs can have information about the past as projective content under negation: *John used to smoke* is projective content under negation of "John stopped smoking." There are many other types of projective contents (e.g., the complement of *know*) under different operators (e.g., questions, modals) identified and discussed in the literature. The problem of explaining how certain inferences can survive entailment-canceling operators is called the *projection problem*.

There are two main approaches to the projection problem. According to the *semantic* approach, projective contents are conventional properties of lexical items (e.g., Frege, 1948; Heim & Kratzer, 1998). According to the *pragmatic* approach, projection can be derived from general conversational principles (e.g., Stalnaker, 1974; Simons, 2001, 2006). How could we capture projection patterns using general conversational principles? To illustrate the reasoning, let us expand the example scenario: Alice and Bob are talking about John, an old friend of Alice's who is visiting her. Bob has never met John before so he knows nothing about him. Bob asks Alice, "Does John smoke?" Alice replies, "John did not stop smoking."[1] Taken literally, Alice's utterance seems under-informative: it can be literally true, regardless of whether or not John smokes. If Alice knows whether John smokes and is cooperative, she would not have said something under-informative. So perhaps her answer is informative after all, but how? Maybe she has taken some additional information for granted, assuming that it is in the *common ground* with Bob. Indeed, if Alice took for granted that John smoked in the past, then, together with "John did not stop smoking," this information would mean that John still smokes, which fully answers the question of whether John smokes now. In other words, assuming that Alice took for granted that John smoked in the past best explains Alice's utterance. If Bob further assumes that if Alice took it for granted then it must be true, then he will arrive at the projected content: John used to smoke.

There are different types of projective content (Tonhauser, Beaver, Roberts, & Simons, 2013). It is possible that they project for different reasons. For change-of-state verbs, there are several reasons why one might prefer a pragmatic approach to projection to a semantic one. First, they systematically show projective behavior. Therefore, a generalization would be missing if their projective contents are lexically-encoded properties that could vary arbitrarily from verb to verb. In contrast, a pragmatic approach could in principle explain why a class of verbs with a similar basic meaning would also have the same projection behavior. Second, projection interacts with the contextual question under discussion, as can be seen from the following example (Geurts, 1995). Imagine that Bob asked Alice: "I notice that John keeps chewing on his pencil. Did he recently stop smoking?" In this context Bob is not interested in whether John is currently a smoker, but

[1] Alice's answer is indirect and complex, and hence would be infelicitous without additional contextual justification. This is predicted by our model.

in whether there was a change from smoking to non-smoking which could explain John's odd behavior. As a result, if Alice were to answer "no (it's just an nervous habit)", Bob would not infer that John used to smoke. Similarly, if a customer asks whether an item on sale has been used by anyone before, a reply "it is not refurbished" would imply that it is brand new. Third, projection is sensitive to prosodic focus: "John did not **stop** smoking" seems to suggest he never smoked. Yet a major obstacle to adopting the pragmatic approach has been the difficulty in formalizing the reasoning and showing that it emerges naturally from conversational principles.

In this paper, we build on and formalize previous ideas of the pragmatic approach. We do so within the Rational Speech-Acts (RSA) framework (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013): a Bayesian approach to language understanding. We extend the previous models by allowing the listener to reason about the facts that the speaker took to be in common ground (along similar lines to Degen, Tessler, and Goodman (2015)). We find that to account for the example scenario we must also make certain assumptions about which facts are plausible common ground and which question is under discussion. With these assumptions and extensions, the model accounts for projection phenomena of change-of-state verbs under negation. It further predicts an interaction between projective behavior and the *question under discussion* (QUD) (Roberts, 2012), suggesting further experimental research to the growing body of literature (e.g., Cummins, Amaral, & Katsos, 2013; Schwarz, 2015)

A Rational Speech-Act (RSA) model

In this section we introduce an extension to Rational Speech-Act (RSA) model (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013) to account for the projection phenomenon of change-of-state verbs under negation, by formalizing the ideas introduced in the previous section. We will continue to use our working example of the conversation between Alice and Bob regarding John's smoking habit.

We consider the following relevant utterances: "John smokes," "John smoked," "John has always smoked," "John stopped smoking," "John started smoking," "John has never smoked," and their negations. In addition, we introduce the null utterance "" (say nothing). The prior probability of an utterance depends on the number of content words (i.e., negation and auxiliaries are excluded) that it has. The shorter an utterance, the higher its prior probability is, as defined in (1).

$$\Pr(u) \propto 2^{-\#\text{content-words}(u)} \tag{1}$$

The meaning/denotation of an utterance is standardly defined as the set of worlds where the utterance is true. We define a world w as a pair. Its first element is whether John smoked in the past and its second element is whether John smokes now. This gives us a set of four possible worlds (the *universe* $U = \{(T,T),(T,F),(F,T),(F,F)\}$). All positive utterances and their denotations are listed in Table 1. In addition, we define that saying nothing is always true, and that the denotation of the negation of an utterance u is $U - [\![u]\!]$.

u	$[\![u]\!]$
"John smokes"	$\{(T,T),(F,T)\}$
"John smoked"	$\{(T,T),(T,F)\}$
"John has always smoked"	$\{(T,T)\}$
"John stopped smoking"	$\{(T,F)\}$
"John started smoking"	$\{(F,T)\}$
"John has never smoked"	$\{(F,F)\}$

Table 1: Positive utterances and their denotations

A *Question Under Discussion* (QUD) is a function Q that takes a possible world as its argument and returns the answer to the question in this world. For example, QUD_{now} is the question "Does John smoke now?" It takes a world and returns its second element, which answers whether John smokes now. Another example is the maximal QUD_{max}, which is the identity function. Intuitively, QUD_{max} is asking which is the actual world. It is maximal in the sense that knowing its answer means knowing the answer to any QUD.

To account for projective behavior, we propose additional components and assumptions to the standard RSA model in the literature (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013; Goodman & Lassiter, 2015). To better illustrate why each of them is necessary and how they contribute to the model's prediction, we will present the model incrementally. We will start with the standard RSA model, point out its problems, motivate a modification, explain the problem it addresses, review the remaining issues, motivate another modification, and so on, until we reach the final model.

Standard RSA model

In the standard RSA model (augmented with QUD as in Goodman and Lassiter (2015)), the literal listener, given an utterance and a QUD, randomly samples a world that is consistent with the utterance, and returns the value of the QUD in that world, as in (2). In this paper we always assume that all worlds are equally likely *a priori*, i.e. $\Pr(w) = 1/4$ for each w.

$$L_0(Q(w) \mid u, Q) \propto \sum_{w' \in [\![u]\!]} \delta_{Q(w)=Q(w')} \cdot \Pr(w') \tag{2}$$

Here δ subscripted with a statement is defined to be 1 if the statement is true and 0 otherwise. For example, if Q is QUD_{max}, which asks for a complete specification of the state of the world, after hearing "John didn't stop smoking," the literal listener will rule out world (T,F), and return the remaining 3 worlds with equal probability.

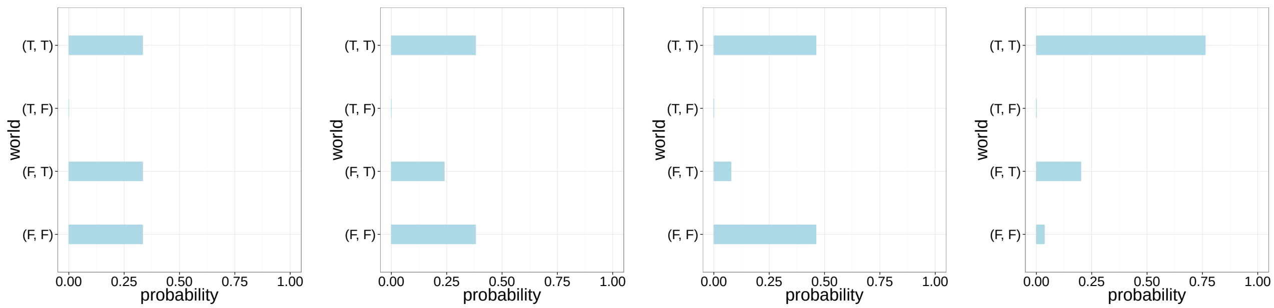
Given the actual world and the QUD, the probability of the speaker's utterance u depends on two factors: the utterance prior and the probability that the utterance will make the literal listener return the correct answer to the QUD, as in (3).

$$S(u|w,Q) \propto \Pr(u) \cdot L_0(Q(w) \mid u, Q)^{\alpha} \tag{3}$$

Here α is a rationality parameter controlling the extent to which the speaker optimizes her utterance to induce the correct answer from the literal listener. When $\alpha \to \infty$, the speaker will always choose utterances that strictly maximize the probability of inducing the right answer. In this paper we set $\alpha = 6$, but the qualitative predictions do not hinge on this specific value.

	Standard (no CG + $\text{QUD}_{\max}$)	Uniform CS + $\text{QUD}_{\max}$	CG prior + $\text{QUD}_{\max}$	CG prior + QUD_{now}
literal	$L_0(Q(w) \mid u,Q) \propto \sum_{w' \in [\![u]\!]} \delta_{Q(w)=Q(w')} \cdot \Pr(w')$	$L_0(Q(w) \mid u,C,Q) \propto \sum_{w' \in C \cap [\![u]\!]} \delta_{Q(w)=Q(w')} \cdot \Pr(w')$		
speaker	$S(u \mid w,Q) \propto \Pr(u) \cdot L_0(Q(w) \mid u,Q)^{\alpha}$	$S(u \mid w,C,Q) \propto \Pr(u) \cdot L_0(Q(w) \mid u,C,Q)^{\alpha}$		
listener	$L(w \mid u,Q) \propto \Pr(w) \cdot S(u \mid w,Q)$	$L(w,C \mid u,Q) \propto \Pr(w) \cdot \Pr(C) \cdot S(u \mid w,C,Q)$		
CG prior	–	$\Pr(C) \propto 1$	$\Pr(C) \propto \sum_{CG \subseteq \text{Obs}} P(CG) \cdot \delta_{C=\cap CG}$	
QUD	$\text{QUD}_{\max}(w) = w$	$\text{QUD}_{\max}(w) = w$	$\text{QUD}_{\max}(w) = w$	$\text{QUD}_{\text{now}}((x,y)) = y$

Table 2: Specifications of four RSA models , with $\Pr(w) \propto 1$ and $\Pr(u) \propto 2^{-\#\text{content-words}(u)}$ for all four models



(a) No CG + $\text{QUD}_{\max}$ (b) Uniform CS + $\text{QUD}_{\max}$ (c) CG prior + $\text{QUD}_{\max}$ (d) CG prior + QUD_{now}

Figure 1: Pragmatic listener after hearing "John did not stop smoking" for each model, with $\alpha = 6$

The pragmatic listener, given the QUD, infers the actual world given the speaker's utterance, using Bayes' rule (4).

$$L(w \mid u,Q) \propto \Pr(w) \cdot S(u \mid w,Q) \quad (4)$$

The standard RSA model is summarized in the first column of Table 2, and the predicted pragmatic listener's distribution over worlds is shown in Figure 1(a). As we can see, the standard RSA model predicts a uniform distribution over the three worlds that are consistent with the literal meaning of "John did not stop smoking". It therefore fails to capture the projective content — the inference that John used to smoke.

The reason for this failure is that "John did not stop smoking" is equally under-informative in any of the three worlds compatible with its literal meaning. For example, suppose the actual world is (T,T). Since the literal listener will return this world with probability only $1/3$ after hearing "John did not stop smoking," the speaker is unlikely to choose this utterance. She is more likely to say "John has always smoked" instead, which will always induce the correct answer. The same holds for the other two worlds (F,T) and (F,F) and therefore the pragmatic listener in the standard RSA model will infer that the three worlds are equally likely.

RSA with common ground

We have seen that one important reason that the standard RSA model fails to capture the projective content of "John did not stop smoking" is that its literal meaning is under-informative when considered in the entire universe U. However, as discussed before, there can be information taken for granted by the speaker and the listener, i.e., the common ground, and an utterance that is under-informative when considered in the entire universe U may nevertheless be informative when evaluated in the common ground. To formalize this observation, we now add common ground to the RSA model.

We first define a related notion. A *context set* C is a non-empty subset of the universe (Stalnaker, 1974). Since we have 4 possible worlds, there are $2^4 - 1 = 15$ different context sets. These context sets are intuitively named. For example, `+past` is the context set that contains (T,T) and (T,F), `+past+now` contains only (T,T), `~+past+now` contains all the worlds except (T,T) ($\sim A$ is defined to be $U - A$, i.e., A's complement), and `change` is the context set that contains (T,F) and (F,T).

A literal listener, given an utterance, the current context set and QUD, randomly samples a world that is consistent with both the utterance and the context set, and returns the value of the QUD in that world, as in (5).

$$L_0(Q(w) \mid u,C,Q) \propto \sum_{w' \in C \cap [\![u]\!]} \delta_{Q(w)=Q(w')} \cdot \Pr(w') \quad (5)$$

For example, given context set `+past` and $\text{QUD}_{\max}$, after hearing "John did not stop smoking," the literal listener will rule out (T,F) because of the utterance's literal meaning, and (F,T) and (F,F) because they are incompatible with the context set. Therefore he will always return (T,T). We can see from this example that an utterance that is under-informative when the entire universe is considered can be informative in some other context sets.

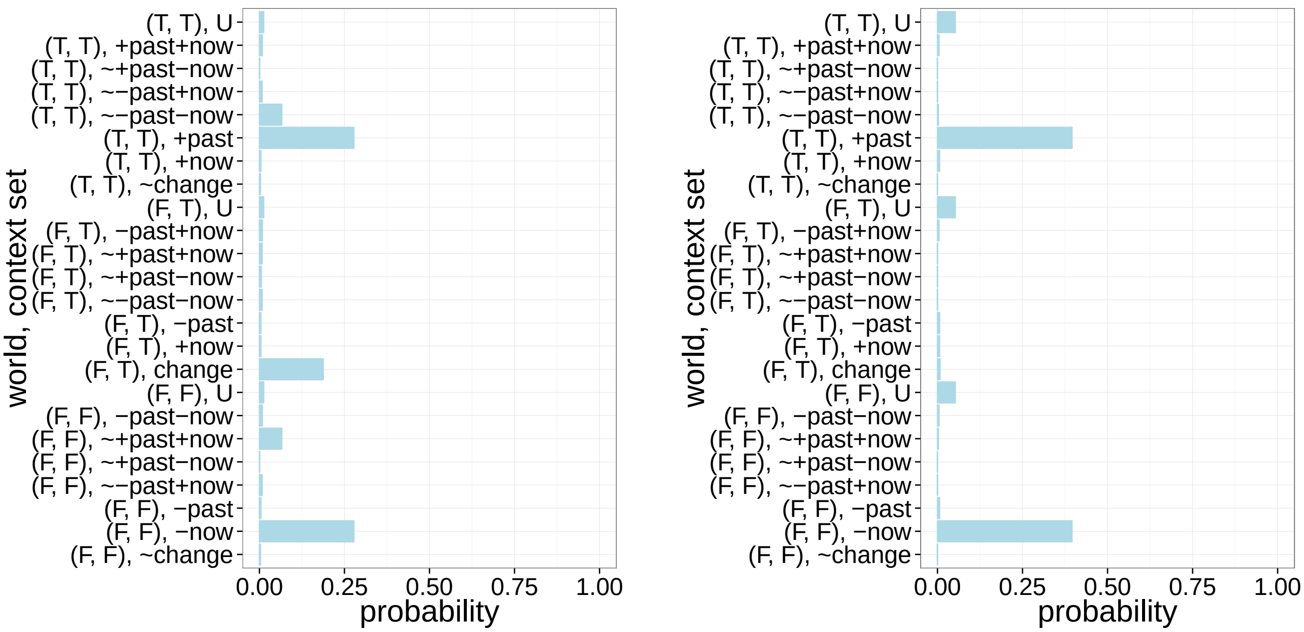
The new speaker model is almost the same as (3), except that it is relativized to the current context set, as in (6).

$$S(u \mid w,C,Q) \propto \Pr(u) \cdot L_0(Q(w) \mid u,C,Q)^{\alpha} \quad (6)$$

Finally, given an utterance and the QUD, the pragmatic listener now jointly infers the real world and the context set the speaker assumes when she produces the utterance.

$$L(w,C \mid u,Q) \propto \Pr(w) \cdot \Pr(C) \cdot S(u \mid w,C,Q) \quad (7)$$

We need to specify a prior distribution $\Pr(C)$ over context sets in (7). We consider two possibilities. First, we consider a uniform distribution over all context sets, i.e., $\Pr(C) \propto 1$.



(a) Uniform CS + QUD_{max}

(b) CG prior + QUD_{max}

Figure 2: Pragmatic listener after hearing "John did not stop smoking," with $\alpha = 6$

Assuming the maximal QUD, the model is summarized in the second column of Table 2 and the pragmatic listener's marginal distribution over worlds is shown in Figure 1(b). We can see that this model predicts that (F,T) is slightly less likely than (T,T) and (F,F), and (T,T) has the same probability as (F,F). This does not capture projection.

The second possibility makes use of the notion of a *common ground* (*CG*) in the pragmatic approach to derive a prior over context sets (Stalnaker, 1974). Intuitively, a common ground represents everything that is taken for granted for conversational purposes. Formally it is a set of propositions (a proposition is a set of worlds), all of which are taken for granted. The context set C, as defined above, can be thought of as the conjunction of all of the propositions that are being taken for granted: $C = \bigcap CG$.

In our example scenario, Alice (the speaker) could reasonably take for granted certain propositions representing plausible observations about John's smoking habits — whether he smoked in the past, and whether he does now. Therefore, assuming that the propositions in the common ground come from observations about John's past and present smoking habits, we can use (8) to naturally define a prior over context sets (henceforth the CG prior).

$$\Pr(C) \propto \sum_{CG \subseteq \text{Obs}} P(CG) \cdot \delta_{C=\cap CG} \qquad (8)$$

Concretely, we assume that each of the observations enters the common ground independently, with probability .4 (meaning that the speaker does not tend to take things for granted). In addition, we add a small amount (5%) of noise to (8), so that every non-empty C has a nonzero prior probability. This model assigns low prior probability to those context sets that cannot be built up via conjunctions of natural observations. One example of such a context set is `change`, the rather complex assumption that John has *either* switched from smoking to not, *or* switched from not smoking to smoking — but we do not know which. In contrast, context sets such as `+past` (i.e., taking for granted that John used to smoke) and `-past-now` (i.e., John did not smoke and does not smoke now) receive higher probabilities because they correspond to observations that the speaker could plausibly have made.

Using the CG prior (the model is summarized in the third column of Table 2), the pragmatic listener's marginal distribution over worlds is shown in Figure 1(c). We can see that this model predicts that world (F,T) is very unlikely, and world (T,T) has the same probability as world (F,F). Although it still does not capture projection because (T,T) is predicted to be as likely as (F,F), the model correctly predicts that (F,T) is unlikely. Therefore we have made some progress.

To better understand how the CG prior improves the model and what the remaining problem is, we plot the pragmatic listener's joint distribution of world and context set in Figure 2.

In Figure 2(a), with a uniform prior over context sets, the pragmatic listener has 3 most likely outcomes: world (T,T) with context set `+past`, world (F,F) with context set `-now`, and world (F,T) with context set `change` (this last outcome is slightly less likely than the first two). This is because in these outcomes, "John did not stop smoking" can fully identify the world given the context set, and hence these outcomes best explain the speaker's utterance. As a result, the marginal distribution over worlds is almost uniform over the 3 worlds.

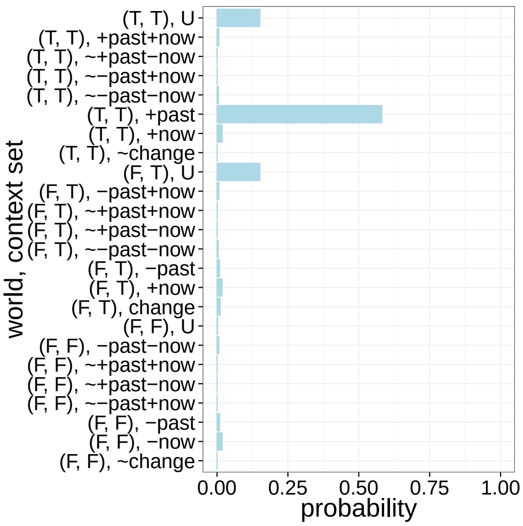


Figure 3: Pragmatic listener after hearing "John did not stop smoking," with $\alpha = 6$, CG prior + QUD_{now}

In contrast, in Figure 2(b), with the CG prior, world (F,T) with context set `change` is no longer a likely outcome, because as noted earlier, the context set `change` is assigned a very low prior. This is why world (F,T) is correctly predicted to be unlikely. Although the CG prior we introduce above is probably over-simplified, the crucial assumption we need is just that not all context sets are equally likely *a priori*, and in particular `change` is a fairly unusual context set and should be assigned a low prior probability, which seems intuitively plausible. As long as this assumption is satisfied, there could be alternative ways to motivate a more realistic prior over context sets without affecting the model's qualitative prediction.

Nevertheless, we can see that world (F,F) with common ground `-now` is still one of the most likely outcomes in Figure 2(b), and hence the marginal probability of (F,F) is the same as (T,T) in Figure 1(c). This is not desirable, but is totally expected from the model: the prior for context set `-now` is the same as for context set `+past`. Therefore, to fully capture projective behavior, we need to further explain why (F,F) with context set `-now` is dispreferred.

Non-maximal QUDs So far, we have been assuming that the QUD is maximal, i.e., the utterance "John did not stop smoking" is chosen to address the question of whether John smoked in the past and whether John smokes now. For this QUD, the RSA model with common ground prior predicts a tie between (T,T) with context set `+past` and (F,F) with context set `-now`.

The maximal QUD is often assumed in applications of RSA models (though see Kao, Bergen, and Goodman (2014)), but in this case there are good reasons to consider non-maximal QUDs. Empirically, as noted in the beginning, projection is sensitive to the QUD. Theoretically, there has been a lot of discussion in the previous literature on the relation between at-issueness and projection (Beaver, 2010; Simons, Tonhauser, Beaver, & Roberts, 2010). Their notion of at-issueness roughly corresponds to QUDs in our model, which may be non-maximal.

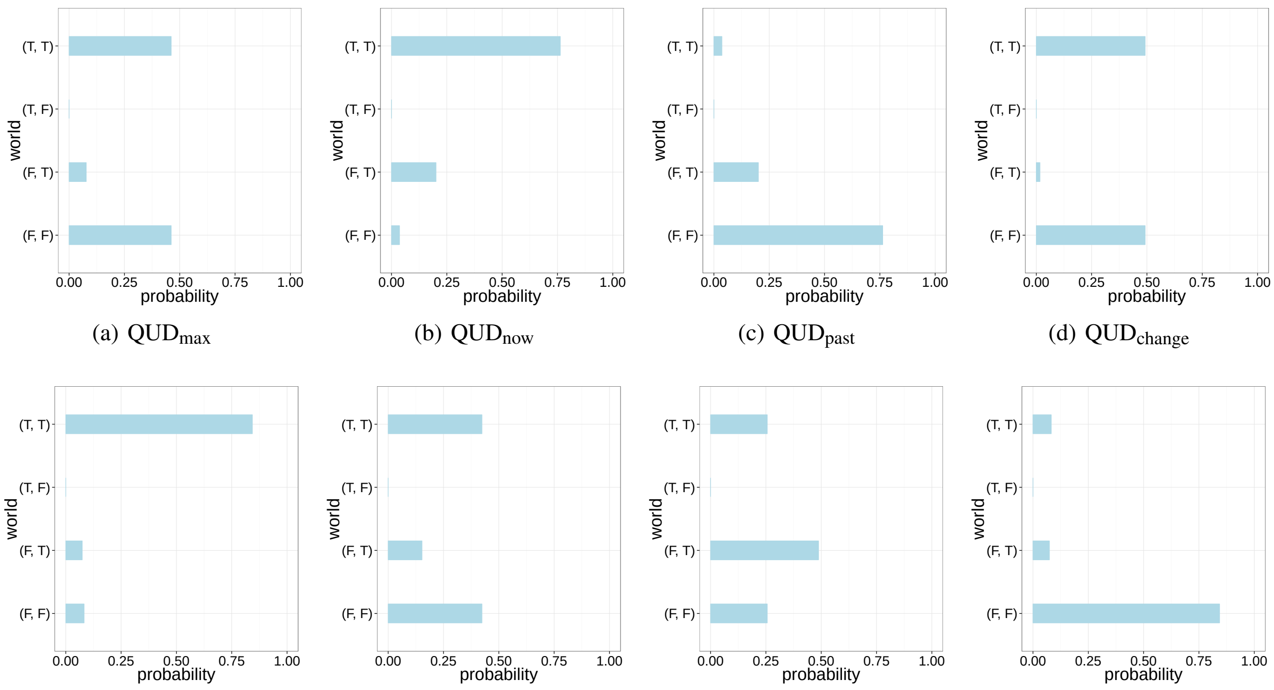
In our running example, Bob explicitly asked about whether John smokes, which means that the QUD is presumably QUD_{now} (i.e., "Does John smoke?"). When we use the previous RSA model with the common ground prior, but replace QUD_{max} with QUD_{now} (summarized in the last column of Table 2), the pragmatic listener's marginal distribution over worlds is shown in Figure 1(d) and the joint distribution of world and context set is in Figure 3. We can see from Figure 3 that (T,T) with context set `+past` is the only most likely outcome, and the world (T,T) is the only most likely world (and its probability increases with a higher α). This is exactly the projection pattern we aim to capture.

To understand why we obtain this result, we note that when the QUD is QUD_{now}, (F,F) with context set `-now` is dispreferred because the context set `-now` already entails the answer to the QUD. That is, it is already known from the context set `-now` that John does not smoke now. This means that the speaker would be maximally informative even if he says nothing. As a result, the speaker would be unlikely to say "John did not stop smoking" when the context set is `-now`, and the pragmatic listener could therefore infer that the context set `-now` is unlikely, which means that (T,T) with context set `+past` is the only winner.

Hence the current model predicts the projective behavior for "John did not stop smoking" in the example scenario, where the QUD is explicitly given by Bob's question. Assuming that people generally care about information about now rather than the past, i.e., the default or most salient QUD is QUD_{now}, the model predicts that the preferred projective content of "John did not stop smoking" without explicit QUD is that John smoked in the past. Note that our speaker model predicts that "John did not stop smoking" is very unlikely to be used to answer QUD_{now}, as there exist simpler utterances "John smokes/does not smoke." This explains the perceived weirdness of Alice's indirect answer to Bob's explicit question.

Other QUDs We have introduced a RSA model with common ground and shown its predictions for QUD_{now} and QUD_{max}. The prediction is sensitive to the QUD—in Figure 4 we show predictions for eight different QUDs. In general, it seems that the model is making plausible predictions. The utterance "John did not stop smoking" implies that John has always smoked if the QUD is whether John has always smoked (Figure 4(e)).[2] It implies that John has never smoked when the QUD is whether John has never smoked (Figure 4(h)). When the QUD is about whether John stopped smoking (Figure 4(f)) or there is a change (Figure 4(d)), the listener will believe that John smoking in the past is about 50% likely (recall Geurts' example described earlier). These results are compatible with Simons et al.'s generalization that only non-at-issue content projects. But further experimental data will be needed to assess whether they are borne out.

[2]Note that QUDs are just partitions and not presuppositional. We use, e.g., *always* to describe the QUDs only as a mnemonic device.



(a) QUD_{max} (b) QUD_{now} (c) QUD_{past} (d) QUD_{change}

(e) QUD_{always} (f) QUD_{stop} (g) QUD_{start} (h) QUD_{never}

Figure 4: Pragmatic listener after hearing "John did not stop smoking" for different QUDs, with $\alpha = 6$ and CG prior

Conclusion and future directions

In this paper, we introduced a probabilistic model in the RSA framework, which analyzes the projective content of change-of-state verbs under negation as the result of the listener using general conversational principles to jointly infer the actual world and the context set that the speaker assumes The model predicts an interaction between projection and the question under discussion, formalizing insights of previous pragmatic approach to projection and providing concrete quantitative predictions that we plan to test experimentally.

Our model is a first step towards a full integration of pragmatic approach to the projection problem into a general probabilistic model of language understanding. We plan to further explore the model and see to what extent it can be generalized to other types of projective content and other entailment-canceling operators, which will help us further understand the division of labor between semantics and pragmatics in the projection phenomena.

References

Beaver, D. (2010). Have you noticed that your belly button lint colour is related to the colour of your clothing. In *Presuppositions and discourse: Essays offered to Hans Kamp* (pp. 65–99). Philadelphia, PA: Elsevier.

Carroll, L. (1866). *Alice's Adventures in Wonderland*. Macmillan.

Cummins, C., Amaral, P., & Katsos, N. (2013). Backgrounding and accommodation of presuppositions: an experimental approach. In E. Chemla, V. Homer, & G. Winterstein (Eds.), *Proceedings of Sinn and Bedeutung 17* (pp. 201–218).

Degen, J., Tessler, M. H., & Goodman, N. D. (2015). Wonky worlds: Listeners revise world knowledge when utterances are odd. In *Proceedings of the thirty-seventh annual conference of the Cognitive Science Society.*

Frank, M. C., & Goodman, N. D. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*(6084), 998. doi: 10.1126/science.1218633

Frege, G. (1948). Sense and reference. *Philosophical Review*, *57*, 209–230. (Originally published in 1892)

Geurts, B. (1995). *Presupposing*. Doctoral dissertation, University of Stuttgart.

Goodman, N. D., & Lassiter, D. (2015). Probabilistic semantics and pragmatics: Uncertainty in language and thought. In S. Lappin & C. Fox (Eds.), *The handbook of contemporary semantic theory, 2nd edition.* Wiley-Blackwell.

Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, *5*(1), 173–184. doi: 10.1111/tops.12007

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Blackwell Oxford.

Kao, J. T., Bergen, L., & Goodman, N. D. (2014). Formalizing the pragmatics of metaphor understanding. In *Proceedings of the thirty-sixth annual conference of the Cognitive Science Society.*

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, *5*(6), 1–69. doi: 10.3765/sp.5.6

Schwarz, F. (Ed.). (2015). *Experimental perspectives on presuppositions.* Switzerland: Springer International Publishing.

Simons, M. (2001). On the conversational basis of some presuppositions. In R. Hastings, B. Jackson, & Z. Zvolenszky (Eds.), *Semantics and linguistic theory (SALT) 11* (pp. 431–448). Ithaca, NY: Cornell University.

Simons, M. (2006). Foundational issues in presupposition. *Philosophy Compass*, *1*(4), 357–372.

Simons, M., Tonhauser, J., Beaver, D., & Roberts, C. (2010). What projects and why. In N. Li & D. Lutz (Eds.), *Semantics and linguistic theory (SALT) 20* (pp. 309–327). Ithaca, NY: Cornell University.

Stalnaker, R. C. (1974). Pragmatic presuppositions. In M. K. Munitz & P. K. Unger (Eds.), *Semantics and philosophy* (pp. 197–214). New York: New York University Press.

Tonhauser, J., Beaver, D., Roberts, C., & Simons, M. (2013). Toward a taxonomy of projective content. *Language*, *89*(1), 66–109. doi: 10.1353/lan.2013.0001