

Guess Who’s Coming (and Who’s Going): Bringing Perspective to the Rational Speech Acts Framework

Carolyn Jane Anderson and Brian Dillon

University of Massachusetts
N408 Integrative Learning Center
650 North Pleasant Street
Amherst, MA 01003, USA
carolynander@umass.edu

Abstract

We present a Rational Speech Acts approach to modeling how conversation participants reason about perspectival expressions. The interpretation of perspectival expressions, such as the motion verbs *come* and *go*, depends on the point-of-view from which they are evaluated. In order to interpret a perspectival expression, the listener must jointly reason about the speaker’s intended message and their choice of perspective. We propose a Bayesian approach to this inference problem and describe an extension of the Rational Speech Acts model that incorporates perspective. We lay out three sets of predictions that this model makes relating to the lexical semantics of *go*, the cost of non-speaker perspectives, and marginal inference over worlds.

1 Introduction

Recent experimental and theoretical work at the semantics/pragmatics interface has highlighted the role of perspective. Perspectival expressions are items whose meaning depends on the point-of-view from which they are interpreted. A wide range of linguistic phenomena appear sensitive to perspective, including spatial and temporal deictic expressions like *on the right* and *tomorrow* (Speas, 2000); pronouns (Loveland, 1984; Wechsler, 2010); motion verbs (Fillmore, 1975; Barlew, 2017); epithets and expressives (Doron, 1991; Harris, 2012); and logophors (Huang and Liu, 2001; Park, 2017).

Perspectival items express two kinds of information: their lexical meaning, and information about

the point-of-view adopted by the speaker. For instance, consider the sentence in (1), containing the perspectival motion verb *come*. *Come* describes motion relative to the location of a perspective holder (Fillmore, 1975). Thus (1) conveys information about the perspective holder’s location in addition to its literal meaning (that Thera is traveling to Northampton).

1. Thera is coming to Northampton in an hour.

In English,¹ motion verbs like *come* and *go* allow at least three kinds of perspective holders: the speaker (e.g. *You are coming to my house*), the listener (e.g. *I am coming to your house*), or the subject of an attitude verb (e.g. *Thera says I am coming to her house*). In order to understand who is located in Northampton, however, the listener must infer who the perspective holder is. How does she do this? That is the central question that this paper addresses.

We propose to model the process of inferring a perspective holder for a perspectival expression using the Rational Speech Acts model (RSA), a framework for pragmatic modeling rooted in Bayesian inference (Frank and Goodman, 2012; Goodman and Stuhlmüller, 2013). We extend the RSA to include the inference required to identify a perspective holder in context. In particular, we model the interpretation of perspectival utterances as inference over the joint probability of a world and a perspective given an utterance. We focus on one particular class of perspectival expressions, perspectival

¹The set of licit perspective holders is a point of cross-linguistic variation: see (Gathercole, 1987) and (Nakazawa, 2007; Nakazawa, 2009) for cross-linguistic work on the topic.

motion verbs, and describe the model’s predictions about some of the open questions related to these verbs, including the lexical semantics of *go* and the existence of a cost for non-speaker perspectives. In what follows we describe the model and articulate its key predictions. We hope that our model and its predictions can guide further exploration of the role of perspective at the semantics-pragmatics interface.

1.1 The semantics of *come* and *go*

Perspectival motion verbs are a good test case for pragmatic modeling of perspective because there is a rich literature on the topic that provides descriptive facts to draw upon and open questions to test. We adopt a view of the lexical semantics of *come* and *go* rooted in the perspectival semantics for motion verbs described in (Barlew, 2017). We use a lexical semantics for *come* consisting of two components: a motion implication, corresponding to the lexical semantics of *come*, and an anchoring implication, corresponding to the perspectival component, expressed in (2) in a simple event semantics.

2. Semantics of *come*:

For any world w , perspective a , destination d , and entity x , $[[Come(x, d)]]^{w,a} = T$ iff

(a) Motion implication:

$[[\exists e. Move(x, e) \wedge Dest(d, e)]]^{w,a} = T$

(b) Anchoring implication:

$\exists y. [[Loc(y, d)]]^{w,a} = T$ and y is a salient perspective-holder with perspective a .

The motion implication is the literal meaning of the sentence: that someone moves to the destination. The anchoring implication expresses that the destination is the location of a perspective-holder.

We use a non-perspectival semantics for *go*: we propose that it has only a motion implication, and no anchoring implication.

3. Semantics of *go*:

For any world w , perspective a , destination d , and entity x , $[[Go(x, d)]]^{w,a} = T$ iff

(a) Motion implication:

$[[\exists e. Move(x, e) \wedge Dest(d, e)]]^{w,a} = T$.

There are a few key points about the meaning of these motion verbs that any model must capture:

First, *come* requires that a salient perspective-holder is located at the destination. This means that *come* cannot be used to describe motion by the perspective-holder. If this were not so, the speaker would always be able to describe their own motion with *come*, regardless of the destination.

4. Context: Matilda and Hildegard are having a conversation in New York.

Matilda: I’ve always wanted to visit Antarctica, but it’s so expensive.

Hildegard: I went / #came there last year. It was so cool!

Second, the lexical semantics of *come* do not directly specify the location of the speaker or the listener, but rather the location of the perspective-holder. Although the speaker of (1) (repeated as (5)) may be the perspective-holder (as illustrated by the speaker-oriented follow-up 5a, the listener also could be (as in the listener-oriented follow-up (5b)).

5. Thera is coming to Northampton in an hour.

(a) I wish you could come hang out too!

(b) I wish I could come hang out too!

This means that the listener must successfully infer the perspective-holder in order to determine the interpretation of any sentence using *come*.

Third, although the lexical semantics of *go* presented here do not directly encode any anchoring implication, in practice *go* seems to imply motion away from the perspective holder. If (6) is interpreted from the speaker’s perspective, then the listener will infer that the speaker is not located in Northampton.

6. Thera is going to Northampton in an hour.

The source of this anti-anchoring implication is up for debate. With the semantics proposed here, this implication is derived via Gricean reasoning on the part of the listener: if the perspective holder was located in Northampton, then the speaker should have used *come*; that she didn’t implies that the perspective holder is not there (Wilkins and Hill, 1995). However, in other theories, *go* does have a perspectival component (Fillmore, 1975; Oshima, 2006).

2 The pragmatics of conversations

2.1 Rational Speech Act

In his foundational work on pragmatics, Lewis (1979) proposed that conversation is a cooperative game between the participants, where the goal is to determine which world the participants are in. Participants work towards this goal by sharing information, which narrows the set of possible worlds that the real world might be. Information shared between the conversation participants is stored in the Common Ground, which can be viewed as the set of accepted propositions, or as the worlds compatible with those propositions.

The Rational Speech Acts model is a framework for pragmatic modeling that extends this picture by proposing that the listener uses a Bayesian inference process to determine what message the speaker is trying to convey (Frank and Goodman, 2012). In this model, the Common Ground contains not just a set of worlds, but also a probability associated with each world, the probability that it is the real world.

At each turn in the conversation, the speaker selects a world from the set of worlds, simulating a new piece of information that the speaker wishes to contribute, and chooses an utterance to express it. Upon hearing the speaker's utterance, the listener must reason about the message that the speaker is trying to convey. The listener assumes that the speaker selects the sentence that maximizes the probability of the observed world.

The listener interprets the sentence in order to update the probability distribution over possible worlds in the Common Ground, calculating the likelihood of each world given the sentence selected, according to their model of how the speaker picks sentences.

The Rational Speech Acts model is therefore recursive. The listener's model of the speaker is called the **literal speaker**: the listener assumes that the speaker estimates the utility of each sentence based on a model of the listener, called the **literal listener**, so called because it does not involve any pragmatic reasoning. The literal listener is what the listener imagines the speaker's model of the listener to be. The listener herself is called the **pragmatic listener**.

The Rational Speech Acts framework has been applied to a range of pragmatic phenomena, including projective content (Qing et al., 2016); polite-

ness (Yoon et al., 2016); scalar implicatures (Potts et al., 2016); and word learning (Smith et al., 2013).

2.2 The standard RSA model

Although a Rational Speech Acts model is potentially infinitely recursive, we limit our discussion to three levels: the literal listener, the literal speaker, and the pragmatic listener.

For the literal listener, the posterior probability of a world given an utterance is simply the literal meaning of the utterance discounted by the prior probability of the world. The meaning of the utterance is its denotation evaluated with respect to the world. Thus, the interpretation function for an utterance is an indicator function returning 1 or 0 depending on whether the utterance is true of the world.

The literal speaker is the listener's mental model of how the speaker selects sentences. In this model, the speaker selects a sentence in proportion to its utility, which is highest for sentences that maximize the posterior probability of the observed world according to the speaker's model of the listener (the literal listener).

The pragmatic listener is the model of the actual listener. The listener reasons about the speaker's message according to their mental model of the speaker (the literal speaker). The posterior probability over a world given the speaker's utterance is proportional to the likelihood of the literal speaker selecting the utterance given the world, discounted by the listener's prior belief in the world.

7. Standard Rational Speech Acts model

(a) Literal listener:

$$L_0(w|m) \propto [[m]]^w p(w)$$

(b) Literal speaker:

$$S_0(m|w) \propto \text{softmax}(L_0(w|m) - \text{Cost}(m))$$

(c) Pragmatic listener:

$$L_1(w|m) \propto S_0(m|w)p(w)$$

where w = world, m = message

One other consideration is usually introduced into the model: the cost of the utterance. This represents the trade-off between informativity and sentence complexity: although the speaker could theoretically select an utterance that exactly isolates the

observed world, as the size of the set of possible worlds increases, the complexity of this utterance also increases. In practice, people often select a simpler, less informative utterance rather than a maximally informative but more complex utterance. The RSA model encodes this as a cost that penalizes the utility for more complex utterances at the level of the literal speaker, for some definition of complexity. A common cost function is the length of the sentence, although syntactic complexity or processing complexity can also be considered.

There are three factors important in defining a RSA model: the set of possible worlds, the set of utterances, and the priors over the worlds. Changes to each of these affect the competition between utterances and the outcome of the inference process.

In general, a uniform probability distribution over the possible worlds is used, simulating the empty Common Ground at the beginning of a conversation. However, in actual conversations, the initial Common Ground may include world knowledge that the speaker and listener have not explicitly mentioned, but are likely to agree upon.

3 The pragmatics of perspective

What about perspectival expressions? In recent work, Roberts (2015) proposes that the Common Ground tracks a set of salient perspectives along with the set of worlds. Perspectival expressions, like *come*, are interpreted not just with respect to a world, but also with respect to a perspective.

In this view, a perspective is a set of *centered worlds*: a set of pairs consisting of a world in which the perspective-holder’s beliefs are true and the spatio-temporal slice of that world that the perspective-holder that the speaker self-identifies with (Stalnaker, 2014). Intuitively, a perspective can be seen as a variable assignment that picks out the individual in each world who the perspective-holder believes is them. A perspectival expression is like an expression with free variable: it cannot be interpreted unless the perspective-holder is known.

If the Common Ground is used to track salient perspectives as well as worlds, then the same inference mechanisms that track the probabilities of worlds in the Common Ground can be extended to track the probabilities of the perspectives. Here we

propose just such an extension.

3.1 A perspectival extension to the RSA

We introduce the Perspectival Rational Speech Acts model, which incorporates perspective into the RSA framework. Building on Roberts (2015), we propose that conversation participants track probabilities over the set of perspectives and use this information to reason about the interpretation of sentences. As in the standard RSA model, the listener infers the probability of the speaker’s message by reasoning about the likelihood of each sentence given the message and the prior probabilities over worlds. For perspectival expressions, however, the listener must also infer the probability of speaker adopting a certain perspective. Therefore, in the perspectival RSA, the listener infers a joint probability: the probability of a paired world and perspective given the utterance. This models the way that the listener extracts information about both the speaker’s perspective and the speaker’s message from a sentence containing perspectival items.

The Perspectival RSA model advances the hypothesis that listeners interpret utterances not with respect to a single perspective, but by considering multiple perspectives simultaneously, each weighted by its probability in context. In this, our model departs from much previous work on the perspective in discourse, which implicitly assumes that listeners interpret utterances with respect to a single perspective (Harris, 2012; Kaiser, 2015; Roberts, 2015).

3.2 The perspectival RSA model

In the Perspectival RSA model, the literal listener reasons about the probability of a world given the utterance and the perspective. That is, in the speaker’s model of the listener, the listener has direct access to the chosen perspective. The posterior probability of the world is proportional to the denotation of the utterance evaluated with respect to the world and the perspective, times the prior probability of the world given the perspective.

8. Perspectival Rational Speech Acts model (simplified definition)

(a) Literal listener:

$$L_0(w|m, a) \propto [[m]]^{w,a} p(w|a)$$

(b) Literal speaker:
 $S_0(m|w, a) \propto \text{softmax}(L_0(w|m, a) - \text{Cost}(m))$

(c) Pragmatic listener:
 $L_1(w, a|m) \propto S_0(m|w, a)p(w, a)$

where w = world, m = message, a = perspective

The literal speaker selects an utterance that maximizes the utility of the observed world according to the speaker’s model of the listener, minus the cost of the message (we return to this below). The speaker observes a world, samples a perspective from the set of perspectives, and selects an utterance accordingly.

The pragmatic listener calculates joint probabilities for worlds and perspectives based on the utterance that the speaker selects. The joint posterior they infer is proportional to the likelihood of the utterance given the world and perspective, according to the listener’s model of the speaker, times the prior joint probability of the world and perspective.

3.3 Independence assumption

Above we showed that the inference in our model relies on the joint prior over world / perspective pairs. However, we make an independence assumption in order to simplify the calculation and assume that the prior probability of a world is independent of the prior probability of the perspective.

9. Perspectival Rational Speech Acts model (intermediate definition)

(a) Literal listener:
 $L_0(w|m, a) \propto [[m]]^{w,a}p(w)$

(b) Literal speaker:
 $S_0(m|w, a) \propto \text{softmax}(L_0(w|m, a) - \text{Cost}(m))$

(c) Pragmatic listener:
 $L_1(w, a|m) \propto S_0(m|w, a)p(w)p(a)$

where w = world, m = message, a = perspective

3.4 The cost of a perspective

In addition to the utterance cost included in the standard RSA, we introduce a perspective cost. The perspective cost function penalizes non-speaker perspectives. We use it to explore the idea that the default perspective is that of the speaker. This idea has been widely discussed in the previous literature

on perspective. In theoretical work, Roberts (2015) posits that in the absence of explicit cues to the contrary, the listener always assumes that the speaker is adopting their own perspective. In experimental work, Harris (2012) and Kaiser (2015) find a strong preference for interpreting expressives, epithets, and other perspectival content from the speaker’s perspective, in the absence of explicit cues otherwise.²

10. Perspectival Rational Speech Acts model (final definition)

(a) Literal listener:
 $L_0(w|m, a) \propto [[m]]^{w,a}p(w)$

(b) Literal speaker:
 $S_0(m|w, a) \propto \text{softmax}(L_0(w|m, a) - \text{Cost}(m) - \text{Cost}(a))$

(c) Pragmatic listener:
 $L_1(w, a|m) \propto S_0(m|w, a)p(w)p(a)$

where w = world, m = message, a = perspective

In Section 5.2, we discuss the impact that this parameter has on the model’s predictions.

3.5 Probabilities over perspectives

One important part of our model is the set of perspective holders that listeners track. In the Roberts (2015) view of perspective, the speaker and listener perspectives are automatically entered into the Common Ground and attitude holder perspectives are entered whenever an attitude verb is used.

Another possibility, however, would be to allow a new perspective to be introduced whenever a new entity possessing a mental state is added to the Common Ground. Further work could test the effect that different methods for introducing perspective holders have on the model.

We adopt a uniform prior over perspective holders for the moment. However, exploring different priors on perspectives might shed light on the topic of perspective maintenance. There is experimental evidence that listeners prefer to posit a coherent perspective over multiple utterances by the same speaker (Harris, 2012). This Maintain Perspective principle might be modeled well by a Dirichlet prior

²Note, however, that these experiments involved written statements where the speaker was explicit, but not the listener, unlike the contexts we discuss below.

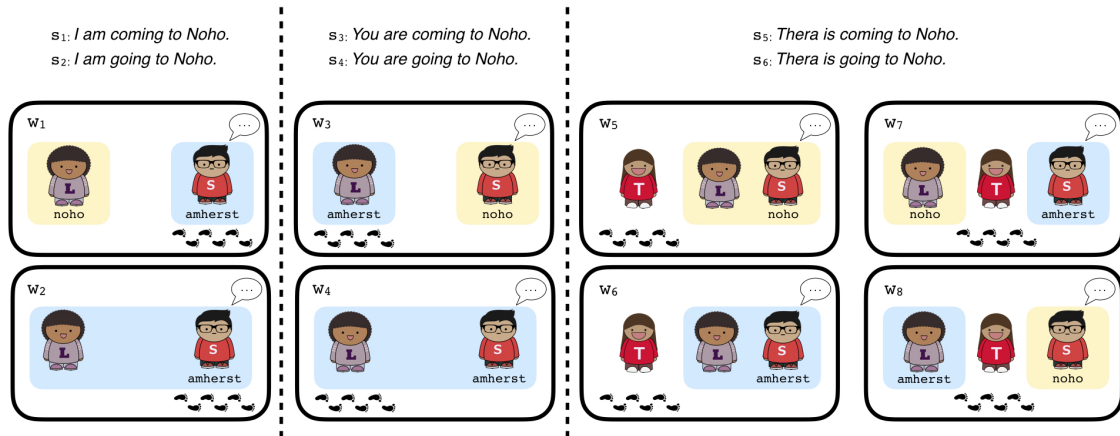


Figure 1: Worlds and utterances.

on perspectives, which would lead to a gradual increase in the probability of perspectives based on how frequently they are adopted by the speaker.

4 An example

To explore the predictions of the perspectival RSA model, we define toy sets of utterances and worlds.

4.1 Perspective holders

We consider a set of worlds with just three entities: Sarah, the speaker; Lydia, the listener, and Thera, a third party who is not involved in the conversation.

Although our model extends to attitude-holder perspectives, we simplify our example by considering only the speaker and listener perspectives. We adopt a uniform prior over the perspective set.

4.2 Worlds

We demonstrate our model with a small set of worlds containing three individuals, Sarah, Lydia, and Thera; and two locations: Northampton (Noho) and Amherst. Since we are not considering Thera in the perspective set, we omit Thera’s location. We include just the worlds in which exactly one person is moving, for a total of eight worlds (Fig.1). We use a uniform prior distribution over the worlds.

4.3 Utterances

We consider two sentence frames: *X is going to Northampton* and *X is coming to Northampton*, where X represents any of the individuals, for a total of 6 possible utterances (Fig. 1). One of the uses of

our model is that it provides a framework for testing the consequences of different lexical semantics for *come* and *go*: we choose the relatively simple denotations presented in Section 1.1 for the sake of illustrating the model, but other semantics for *come* and *go* could be substituted instead.

5 Model predictions

We implemented our model in the WebPPL probabilistic programming language (Goodman and Stuhlmüller, 2014). Using the worlds, perspectives, and utterances described above, we ran simulations with various parameter settings to generate predictions about the probabilities of world/perspective pairs inferred by the listener. For each result reported below, we ran 100,000 iterations of Markov Chain Monte Carlo sampling.

Our key results are: (1) perspectival interpretations of *go* can arise through pragmatic competition even if the lexical semantics of *go* are not perspectival; (2) the likelihood of non-speaker perspectives increases proportionally as the perspective cost is decreased; and (3) listeners should favor worlds that are consistent with multiple perspectives.

5.1 Competition between *come* and *go*

The first result of our model is that perspectival usage of *go* can arise through pragmatic competition with *come* even if the lexical semantics of *go* do not reference perspective. Wilkins and Hill (1995) posit that at least in some languages, the interpretation of *go* as ‘movement away from the perspective holder’

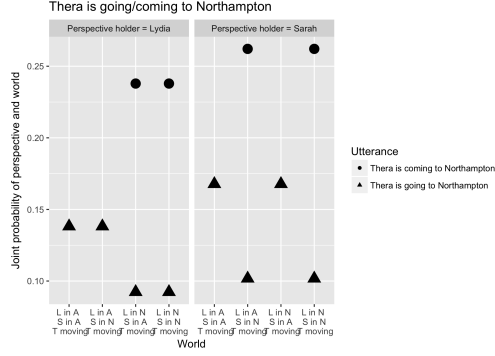


Figure 2: Model predictions for *Thera is going to Northampton* and *Thera is coming to Northampton*, speaker cost = 0.5

is not lexically encoded, but arises instead through pragmatic reasoning. Our model verifies that the perspectival interpretation of *go* could indeed arise solely through pragmatic competition with *come*. As shown in Fig. 2, our pragmatic listener infers that for sentences with *go*, the perspectival holder is unlikely to be located at the destination of motion—even though the lexical semantics of *go* we adopted are not perspectival.

For the *go* sentences, world / perspective pairs where *come* is a viable alternative, such as when the perspective holder is the listener and the listener is at the destination of motion, are less likely than ones where *come* is not a possible alternative.

Thus, even without a perspectival lexical semantics, *go* acquires an interpretation that the perspective holder is unlikely to be at the destination.

5.2 The cost of perspective shift

Another set of predictions relates to the perspective cost parameter. As the cost for non-speaker perspectives increases, the likelihood of a non-speaker perspective decreases. Because our cost function only assigns cost to non-speaker perspectives, with a uniform prior over perspectives, a non-speaker perspective will only be more likely than the speaker perspective when the speaker perspective is excluded by the lexical semantics of the sentence. In the set of utterances we use above, this occurs with only one sentence, *I am coming to Northampton*, where the speaker perspective is excluded because the speaker is moving. For this sentence, the listener will infer with probability 1.0 that the perspective holder is the listener (Lydia) and the world is the one in which

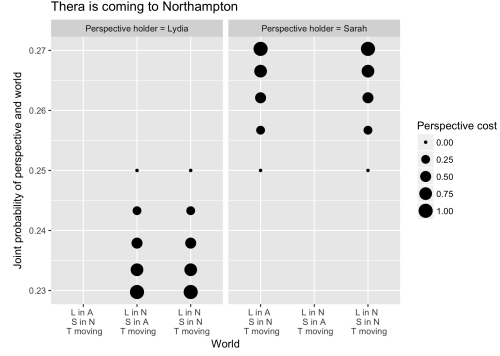


Figure 3: Model predictions for *Thera is coming to Northampton* at varying levels of speaker cost.

Lydia is in Northampton and Sarah is moving.

Although decreasing in probability as perspective cost increases, the listener perspective remains possible for all sentences but *You are coming to Northampton*, where the listener perspective is eliminated because she is in motion. In this case, the speaker perspective is inferred with probability 1.0.

For a third-party mover, the listener perspective is as likely as the speaker perspective only when the perspective cost for non-speakers is set to zero. As the speaker cost increases, the listener perspective becomes proportionally less likely (Fig. 3).

For the *go* sentences, both perspectives are always valid possibilities, since *go* does not have a perspectival component. For these sentences, the likelihood of the speaker perspective increases proportionally as the speaker cost is increased (Fig. 4).

Although the perspective cost parameter can be used to explore whether there is a processing cost for adopting a non-speaker perspective, since the cost is part of the listener’s model of the speaker, it can also be used to explore the listener’s beliefs about the speaker. If the listener has reason to believe that a particular speaker is more or less likely to adopt a non-speaker perspective, the listener could adjust the perspective cost parameter accordingly. This might occur in situations where the speaker has higher rank than the listener; or situations where the listener thinks the speaker has limited knowledge of the listener’s location.

5.3 Marginal inference

As mentioned above, one novel claim of our model is that listeners consider multiple perspectives at

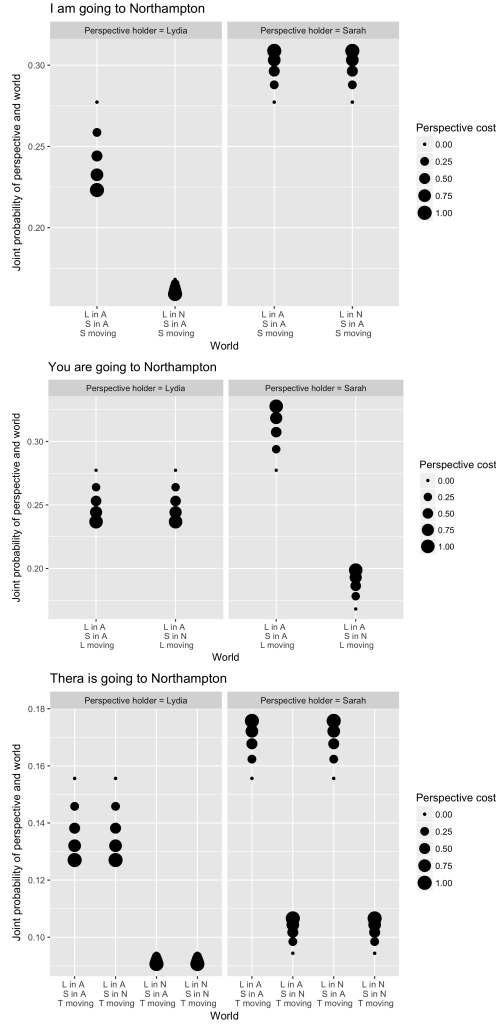


Figure 4: Model predictions for *go* sentences at varying levels of speaker cost.

once when they interpret utterances, rather than assuming that the perspective is that of the speaker, as in Roberts (2015), or the most recently used perspective, as Harris’s perspective maintenance principle suggests (2012).

The consideration of multiple possible perspectives leads to a different marginal distribution over worlds than if the listener assumed one particular perspective. To see how this works, consider the model’s predictions for *Thera is coming to Northampton* when speaker cost is 0.5. The posterior probabilities are shown in Fig. 5.

The marginal distribution strongly favors the world in which both the speaker and the addressee are located at the destination of motion. Although

	Speaker	Listener	Marginal
Both in Noho:	0.26	0.24	0.5
Neither in Noho:	0	0	0
Listener in Noho:	0	0.24	0.24
Speaker in Noho:	0.26	0	0.26

Figure 5: Non-zero posterior probabilities for *Thera is coming to Northampton*, speaker cost = 0.5

we use the posterior distribution in Fig. 5 as an example, this is true regardless of how high the speaker cost is, so long as there is a non-zero chance that the listener is the perspective holder.

On the other hand, if the listener interprets the sentence assuming that the speaker is by default the perspective holder, she would have no reason to prefer the world in which both the speaker and listener are at the destination over the world in which just the speaker is there. All worlds in which the speaker is located at the destination would be equally likely.

In this way, the ‘multiple perspectives’ hypothesis predicts that listeners will prefer to interpret *Thera is coming to Northampton* as conveying a message that is true for all salient perspectives; models that assume a single perspective for interpretation of an utterance do not. We plan to test this in an experiment where speakers are presented with sentences like *Thera is coming to Northampton* and are asked to choose the most likely scenario from a set of visual stimuli depicting the possible worlds. If listeners perform the kind of pragmatic inference that we propose, they should select the world in which both possible perspective holders are at the destination.

6 Conclusion

We have outlined an extension to the Rational Speech Acts model in which the listener jointly reasons about the likelihood of the speaker’s utterance and the adopted perspective. This model provides a framework for exploring the effects of various cost functions, priors, and lexical semantics for perspectival items. We have described some of the predictions that our model makes for one kind of perspectival expression, perspectival motion verbs; we hope the model that we have outlined can serve as a useful framework for generating testable predictions about other kinds of perspectival expressions as well.

References

- Jefferson Barlew. 2017. *The semantics and pragmatics of perspectival expressions in English and Bulu: The case of deictic motion verbs*. Dissertation, Ohio State University.
- Edit Doron. 1991. Point of view as a factor of content. *Proceedings of Semantics and Linguistic Theory Conference (SALT)*.
- Charles J. Fillmore. 1975. *Deixis*.
- Michael C. Frank and Noah D. Goodman. 2012. Predicting pragmatic reasoning in language games. *Science*, 336:998–998.
- Virgina Gathercole. 1987. Towards a universal for deictic verbs of motion. *Kansas Working Papers in Linguistics*, 3.
- Noah D. Goodman and Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5:173–184.
- Noah D. Goodman and Andreas Stuhlmüller. 2014. *The Design and Implementation of Probabilistic Programming Languages*. <http://dippl.org>.
- Jesse Harris. 2012. *Processing Perspectives*. Dissertation, University of Massachusetts, Amherst.
- C. T. James Huang and C. S. Luther Liu. 2001. Logophoricity, attitudes, and *ziji* at the interface. *Long-distance reflexives*, 33:141–195.
- Elsi Kaiser. 2015. Perspective-shifting and free indirect discourse: Experimental investigations.
- David Lewis. 1979. Score-keeping in a language game. *Semantics from a Different Point of View*.
- Katherine A. Loveland. 1984. Learning about points of view: spatial perspective and the acquisition of ‘i/you’. *Journal of Child Language*, 11:535–556.
- Tsuneko Nakazawa. 2007. A typology of the ground of deictic motion verbs as path-conflating verbs: the speaker, addressee, and beyond. *Poznan Studies in Contemporary Linguistics*, 43.
- Tsuneko Nakazawa. 2009. A typology of the ground of deictic motion verbs as path-conflating verbs: the entailment of arrival and the deictic center. *Poznan Studies in Contemporary Linguistics*, 45.
- David Y. Oshima. 2006. Motion deixis, indexicality, and presupposition. *Proceedings of Semantics and Linguistic Theory (SALT)*, 16.
- Yangsook Park. 2017. *Attitudes dese and logophoricity*. Dissertation, University of Massachusetts, Amherst.
- Christopher Potts, Daniel Lassiter, Roger Levy, and Michael C. Frank. 2016. Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. *Journal of Semantics*, 33.
- Ciyang Qing, Noah Goodman, and Daniel Lassiter. 2016. A rational speech-act model of projective content. *Proceedings of the Cognitive Science Society*, 38.
- Craige Roberts. 2015. Indexicality: *de se* semantics and pragmatics.
- Nathaniel J. Smith, 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 (NIPS)*.
- Margaret Speas. 2000. Person and point of view in navajo: direct discourse complements. *Proceedings of West Coast Conference on Formal Linguistics*.
- Robert C. Stalnaker. 2014. *Context*. Oxford University Press.
- Stephen Wechsler. 2010. What ‘you’ and ‘i’ mean to each other: person indexicals, self-ascription, and theory of mind. *Language*, 86:332–365.
- David P. Wilkins and Deborah Hill. 1995. When ‘go’ means ‘come’: Questioning the basicness of basic motion verbs. *Cognitive Linguistics*, 6:209–259.
- Erica J. Yoon, Michael Henry Tessler, Noah D. Goodman, and Michael C. Frank. 2016. Talking with tact: Polite language as a balance between kindness and informativity. *Proceedings of the Cognitive Science Society*.