Games in Linguistics[1]

Nicholas ASHER — *CNRS-IRIT, Toulouse*
Julie HUNTER — *Universitat Pompeu Fabra, Barcelona and Université Paul Sabatier, Toulouse*
Soumya PAUL — *Université Paul Sabatier, Toulouse*

Abstract. In this paper we set out three consequences of a game-theoretic model for conversation, Message Exchange (ME) Games (Asher et al., 2016), which we think are of linguistic interest. We develop a notion of conversational success, explain subjectivity and bias in interpretation using concepts from epistemic game theory, and characterize the strategic usefulness of using so called expressions of "not at issue" content using ME games.

Keywords: epistemic games, conversation, discourse, conversational success, at-issue content, subjectivity

1. Introduction

The philosopher Grice long ago popularized the idea that conversation is a rational activity (Grice, 1975), yet curiously, efforts to apply philosophical and economic analyses of rationality and rational strategizing to linguistic phenomena have been sporadic and very restricted in their aims. Much of this work, including van Rooij (2004); Franke et al. (2012); Franke (2008) and Asher and Lascarides (2013), has been directed to the justification of Gricean maxims of conversation, often with the further aim of computing scalar implicatures. As argued in Asher and Lascarides (2013), the focus on Gricean maxims is largely misplaced: they are not in and of themselves an interesting *linguistic phenomenon*; they are an informal, and somewhat inchoate, description of what more formal models of rational interaction predict. Such formal models, when coupled with a well-developed theory of discourse structure and interpretation, have a much broader range of application to linguistic phenomena. In this paper, we argue that they play a crucial role in the analysis of three particular phenomena: evaluations of conversational success, not-at-issue/at-issue notions of content, and the subjectivity of interpretation. As we will make clear in our analysis of the subjectivity of interpretation, our formal model of rational strategizing affects how we structure and interpret a conversation. As some of us have argued at length that discourse structure affects many dynamic semantic phenomena (temporal structure, the interpretation of anaphora and ellipsis) as well as discourse content as a whole, these models thus have a general importance for understanding content in all its manifestations.

Mathematics, theoretical computer science and economics have produced a rich and pertinent body of work on which to draw in building a model of rational behavior. Conversations, for example, have a natural analysis as games. They typically involve at least two agents, each with their own interests and goals. These goals may be compatible or they may be in conflict, but in either case, one agent's successfully achieving her conversational goals will typically depend upon her taking her interlocutors' goals and interests into account. In cooperative conversations, in which agents' goals are completely aligned, conversational partners typically still need to coordinate actions, even linguistic actions. In strategic or non-cooperative conversations, in which participants have opposing interests concerning the outcome of the conversation, the ne-

[1]We thank ERC grant 269427 for research support.

cessity to consider the opponents' aims and actions is almost always even more important. A debate between two political candidates is an instance: each candidate has a certain number of points to convey to the audience, and each wants to promote her own position and damage her opponent's or opponents'. To achieve these goals, each participant typically needs to plan for anticipated responses from the other.

Our paper is organized into three main parts. First, we look at an application of a game theoretic model to the notion of conversational success and provide an abbreviated description of the technical details of the model, which we call *epistemic message exchange games*. We then show how the model sheds insight on the subjectivity of interpretation. In the third section, we apply the game theoretic model to an analysis of different types of content, in particular the distinction between what linguists call "at-issue" and "not-at-issue" content.

2. Conversational success

While linguists are accustomed to semantic evaluations in terms of truth and satisfaction alone, ordinary people evaluate their conversational contributions and those of others more generally in terms of what we call *success*. Did the agent achieve her conversational goals with her contributions or not? Conversational success thus has to do with the goals a conversationalist has. What, then, are conversational goals? One possibility is to identify a conversational goal extensionally as the set of conversations that are successful from the point of view of the speaker. Conversational goals are then defined as subsets of the set of all possible conversations that exclude those conversations that do not go well. Sometimes a conversation will count as successful in virtue of containing a particular verbal string, as illustrated in (1).

(1) a. EPA administrator : May I look inside the containment structure?
b. Ghostbuster (Bill Murray): You didn't say the magic word.
c. EPA administrator: Please, may I look inside the containment structure?
d. Ghostbuster (Murray): No. (from *Ghostbusters*)

At least one of the goals of Murray's character is simply to have the EPA administrator prefix his request with the word *please*. If we define this goal extensionally, we end up with the set of all conversations in which that string follows (1a) and (1b). Most conversational goals, however, are not defined by particular strings. What matters are the commitments to conversational contents that the interlocutors ultimately adopt. An evaluation of conversational success therefore typically has ties to a conversation's content or its ordinary semantic evaluation. Importantly, this does not mean that the content of a conversation must be *true* or accurate; a conversational contribution may be successful in persuading an interlocutor to do something, for example, even if the contribution is inaccurate or false. Certain 2016 US Presidential campaigns provide ample evidence of this possibility.

Content related goals can be tied to particular discourse moves. Following Asher and Lascarides (2003), asking a non-rhetorical question, for instance, indicates that the speaker has the conversational goal of obtaining an answer from her interlocutor or interlocutors. A simple conversational goal for an assertion is typically to have one's interlocutors agree or at least not

openly object to it. Asher and Lascarides (2003) call such goals *speech act related goals*.

Conversational goals, however, can also global, general properties of a conversation that guide a large stretch of discourse or even the conversation as a whole. Consider, for instance, a prosecutor who either wants a witness to commit to some issue or wants to demonstrate before a jury that the interlocutor is evading or refuses to answer the question. Success may require several discourse moves (and may never be achieved at all). As a real life example, consider the following exchange between CNN's Jake Tapper and Mike Pence, who was the US Vice-President elect at the time.[2] Tapper asks Pence if he was aware that the transition team for Trump's presidency had put in for a security clearance for Michael Flynn Jr., the highly controversial son of Trump's choice for National Security Advisor. Pence repeatedly dodges the direct yes or no question, forcing Tapper to point out why all of Pence's attempts to deflect the question were not answers. While Tapper never succeeds in getting a direct answer from Pence, his eventual success in extracting at least a strong implicature that Pence was aware of the demand for clearance required a series of arguments pointing out why each attempt at diversion by Pence was just that.

More often than not, the goal of a particular conversation such as an interview or a debate will be a combination of simpler conversational goals in some temporal logic like Linear Temporal Logic (LTL) (Lamport, 1980), which includes the temporal operators $\Diamond$ for *eventually* and $\Box$ for *always*, as well as operators for the temporal relations *since* and *until*. A reporter, for example, might have the complex goal of eventually getting a satisfactory answer to each of her individual questions to her interlocutor, and the goal of getting an answer to a question Q, as we have seen, typically has the general form: until an answer to Q is produced, show that no answer to Q has been given and then repeat Q. A more complex goal, which might be adopted by a participant in a political debate, is to reply to every attack on her and to land more attacks on her opponent than he lands on her. This goal is not expressible in LTL, but is in the framework we develop below.[3] Asher et al. (2016) provide many examples of such goals and show how these goals may differ in complexity. They also show how to link goals to strategies for achieving them.

Games provide natural structures within which to investigate the success of sequences of linguistic actions. Game theory evaluates actions in terms of utility, and the simple Boolean case of winning conditions we have alluded to above is an instance of a utility function. Signalling games (Lewis, 1969; Spence, 1973), which have been very popular in linguistics, are not appropriate for the task, however. Signalling games are designed to tackle a different aspect of language, namely, the coordination on linguistic content in reflective equilibrium (see Lewis's account of the emergence of linguistic conventions). By contrast, we are interested in evaluations of conversational success, even—and especially—in cases where interests of the conversationalists are opposed. In such cases the meanings of messages in the context of signalling games is problematic to say the least; Crawford and Sobel (1982) show that in cases of opposing interests messages cease to have content in reflective equilibrium.

[2] 'Pence pressed on clearance for Flynn's son,' *The Lead*. The full exchange can be viewed here: http://edition.cnn.com/videos/politics/2016/12/06/mike-pence-trump-flynn-jr-transition-lead-bts.cnn

[3] See Asher et al. (2016) for details.

Our game theoretic model is different from signalling games in several respects. For one, we will take messages to have an exogenously given meaning that determines how one conversationalist responds to the messages of another. This allows us to avoid the problems of message interpretation in signaling games where the interests of the players are opposed. Another way in which our model differs from signaling games is that while signaling games are typically "one shot", a good model of conversational goals, as suggested in the examples above and argued in detail in Asher et al. (2016), should require agents to strategize about conversations as open-ended sequences of moves with no set end. In the exchange between Jake Tapper and Mike Pence, described above, Tapper had the goal of eventually getting an answer to his question about Michael Flynn Jr., and he didn't stop his line of questioning until he got at least an implied answer; he had to be prepared for an open-ended set of moves by Pence designed to avoid the question. In general, conversational agents must plan for any number of moves by their opponents to try to frustrate or to prevent them from achieving their goals. In fact, we can put no *a priori* upper bound on the number of moves that accomplishing this goal might require, and thus a game theoretic framework for conversation must countenance a potentially infinite sequence of exchanges of messages between conversational participants.

To model these aspects of strategic conversations, Asher et al. (2016) developed a game theoretic framework of *Message Exchange* or ME games. The intuitive idea behind an ME game is that a conversation is a sequence, either finite or infinite, of turns. In each turn, one of the players 'speaks' or plays a sequence of moves, and each sequence of moves itself describes a discourse structure in the sense of SDRT (*Segmented Discourse Representation Theory*; Asher and Lascarides, 2003) that extends the discourse structure built up from previous turns. More precisely, the vocabulary V of an ME game contains a set of discourse unit labels $\mathsf{DU} = \{\pi, \pi_0, \pi_1, \ldots\}$, a set of formulas from a language for dynamic semantics that serve to describe the contents of the basic units (where '$\pi\colon \phi$' means that the formula ϕ describes the contents of the discourse unit π), and a set of discourse relation symbols $\mathscr{R}$ relating discourse constituents from the different moves made so far in the game.

Turns in ME games are relativized to players. In the case of conversations, it is essential to keep track of "who says what"; Tapper saying that Pence had knowledge of Flynn Jr.'s past is not the same as Pence himself admitting to having had this knowledge. To model this, each player i is assigned a copy V_i of the vocabulary V of SDRT moves, which is simply given as $V_i = V \times \{i\}$. Thus when Player i plays $u \in V$, it is noted as (u,i). Conversations correspond to plays of ME games which are finite or infinite sequences over $(V_0 \cup V_1)$, noted as $(V_0 \cup V_1)^\infty$, for a game with two players, 0 and 1 (for details see Asher et al., 2016; Asher and Paul, 2017).

Given that we have defined a conversation as a sequence or element of $(V_0 \cup V_1)^\infty$, a conversational goal will be a subset of $(V_0 \cup V_1)^\infty$. But who determines what the goals are in a given conversation? That is, who determines which subsets of $(V_0 \cup V_1)^\infty$ represent success? Speakers presumably have their internal goals, but it is not necessarily those that determine conversational success. Each person who is involved in or who witnesses a conversation has her own ideas about what the winning conditions of the participants are or should be.

To this end, Asher et al. (2016) introduced a crucial component for analyzing conversations

that they called the *Jury*. The Jury determines the conversational goals of the participants. The Jury is itself an abstract decision rule determining winning conditions, but it can be instantiated with the conversationalists themselves, or with a third party that evaluates the conversation. In strategic settings, taking conversational partners as evaluators can lead to trouble, as each will be tempted to declare him- or herself the winner. In many settings such as political debates, the natural Jury to consider is the actual audience, some segment of the population who witnessed the debate or the whole set of participants, or yet some other body like the editorial board of a newspaper.

Asher et al. (2016) consider only an impartial Jury, who also enforces constraints like the consistency of a player's contributions or the constraint that a player respond to questions or other moves of the other players. For most conversations, however, there will be many Juries. These Juries may disagree with each other about winning conditions, and some may have a very biased take on a conversation. Nevertheless, given what we have said so far, we will need to evaluate conversational success relative to a particular Jury. Accordingly, ME games pair the space of possible conversations $(V_0 \cup V_1)^\infty$ with a Jury $\mathscr{J}$.

Evaluating conversational success is to a certain extent a subjective or relative matter, since it depends on the conversational goals assigned to players by the Jury. Different juries may disagree as to what the conversational goals should be, and a group of people or even a single person may be undecided as to what sort of Jury she is. Still, once a goal is set as a subset of $(V_0 \cup V_1)^\infty$, it is an objective matter as to whether the conversation meets this goal or does not; either the conversational play is an element of the subset designating a conversationalist's winning condition or it is not.

A final point concerns the evaluation of conversational success. When does a Jury decide a conversationalist has met the winning condition assigned to her? While for Asher et al. (2016) a Jury must survey an entire conversational string and all its continuations, which may be infinite, Asher and Paul (2016) argue that this misses the fact that an actual Jury evaluates in a dynamic fashion, after each turn by one of the conversational participants. The idea is to represent a winning condition by a scoring function over players' conversational turns. A turn can be rated more or less good with respect to the winning conditions the Jury has in mind, or more or less disastrous. If the Jury has a scoring function with a discounting parameter that lowers the score for turns later in the conversation, it will always being able to confidently pick a winner in a 0 sum game within a finite amount of time. Asher and Paul (2016) illustrate how such a scoring function works on a snippet from one of the debates between candidates for the Republican Presidential candidate of 2016.

So what is the import of ME games for linguistics? The field of pragmatics has always been concerned with the use of language, and ME games yield a formal pragmatic framework that provides principled reasons for why and how we use language to attain conversational goals. Hearkening back to the concerns of traditional rhetoricians, our account answers the question: what is a reasonable scoring function and how should a conversationalist attempt to maximize her score? To answer this question, the framework of ME games replaces semantic evaluation in terms of truth at a world with pragmatic evaluation in terms of conversational success relative to

a Jury. The underlying structure of game theory also generates a notion of logical consequence for conversational success, a very rough approximation of which would be LTL's notion of consequence, which works for simple goals, though the nature of the full consequence relation is as far we know unexplored. Finally, the framework predicts whether there is a winning conversational strategy for a given goal and what it would look like in linguistic terms.

3. The subjectivity of interpretation

In this section, we explore the nature of a Jury's scoring functions, and in particular how subjectivity and bias naturally influence them. One of the astounding facts about conversations is that people who participate in or merely observe them can come away with dramatically different interpretations of what was said, even though not everything goes. Different Juries can disagree about what was said, what was implied, and about who was successful. Consider, for instance, the one-line retort by Presidential candidate Trump to Presidential candidate Clinton during the third US Presidential debate in 2016:

(2) Such a nasty woman.

The literal meaning and the conventional implicatures of this remark in context are clear; with (2), Trump committed himself to a negative assessment of Clinton. However, people perceived the role of (2) in achieving an agent's conversational goals very differently. One Jury, an appreciable segment of the American population, found this remark totally out of place in a Presidential debate. Another one, Trump's base, found the comment appropriate and would have assigned it a high score. Each of these interpretations depends on how the Jury assigns winning conditions to the players, which includes constraints under which, the Jury judges, conversations should be conducted.

Bias and subjective beliefs will also influence the way that a Jury interprets the very structure of a discourse. Consider the following excerpt, discussed at length in Asher et al. (2016), from a press conference by Senator Coleman's spokesman Sheehan. Senator Coleman was running for re-election as a US senator from Minnesota in the 2008 election.

(3)
a. **Reporter:** On a different subject is there a reason that the Senator won't say whether or not someone else bought some suits for him?
b. **Sheehan:** Rachel, the Senator has reported every gift he has ever received.
c. **Reporter:** That wasn't my question, Cullen.
d. **Sheehan:** (i) The Senator has reported every gift he has ever received. (ii) We are not going to respond to unnamed sources on a blog.
e. **Reporter:** (i) So Senator Coleman's friend has not bought these suits for him? (ii) Is that correct?
f. **Sheehan:** The Senator has reported every gift he has ever received.

Sheehan continues to repeat, *The Senator has reported every gift he has ever received* seven more times in two minutes to every follow up question by the reporter corps.[4]

[4] See `http://www.youtube.com/watch?v=VySnpLoaUrI`.

While many of the contributions by Sheehan and the reporter corps have a clear and uncontroversial meaning and discourse function, some contributions are open to interpretation. For instance, how are we to interpret the response α by Sheehan in (3b), (3d.i), and (3f)? ME games provide an insightful answer. To formulate the above exchange as an ME game, we first fix the players and the Jury. We can assume that there are two active players: the reporter corps (R) and spokesman Sheehan (S). We will also consider two Juries, each of which interprets the exchange differently. The first, Jury 1, starts out with a presumption of full disclosure and honesty from S; Jury 2 is a biased Jury that is disposed to believe whatever the spokesman says, because, for example, Jury 2 and S are from the same political party. We'll see below that Juries 1 and 2 will arrive at different interpretations of the structure and content of (3), and in particular, of S's repeated response.

To model the subjectivity of conversational interpretation, we must clarify what elements of a discourse structure are influenced by subjective interpretation. Asher et al. (2016) and Asher and Paul (2016) assume no ambiguity in the discourse moves made by players in an ME game (though for a look at a prior treatment of ambiguity in the ME setting, see Venant and Asher, 2015). However, at least at a first pass and assuming a perfect communication channel, it is the ambiguous moves that are up for interpretation, so we will need to countenance some ambiguity. At the same time, not everything is up for interpretation—an emphatic *no* to a polar question doesn't mean *yes*.

We will assume that the grammar, including syntax and lexical and compositional semantics, delivers an unambiguous core or, following Asher and Lascarides (2003), an *underspecified logical form* (ULF) for a discourse. We will also distinguish between *plays*, the objective components of discourse moves that are uncontroversially part of the speakers' public commitments, and *histories*. A play in our ME games is a ULF, and given our assumption about exogenously given meaning, we will assume this ULF and its interpretation are common knowledge of the players and the Jury. Of course, a ULF typically involves underspecified elements (whose semantics we can specify via existential quantifications over variables standing for the elements that require specification) that are specified via reasoning that depends on a variety of subjective sources. For (3), the observed play, call it play ρ, is a representation for (a-f), in which each contribution has its normal compositional semantics and some of the uncontroversial discourse connections, like the fact that (c) corrects the discourse connection between (a) and (b), are made explicit. Other relations are left underspecified, including the relation between (a) and (b), (c) and (d.i), and (e) and (f).

(4) provides the ULF for (3), where each π_n labels an SDRT formula, $Sel_{\mathbb{R}}$ is a selection function over discourse relations that signals the presence of underspecified relations, and Sel_{Π} is a selection function over discourse units that marks the presence of underspecified arguments.

(4) $\pi_2 : (\pi_0 : 3a \wedge \pi_1 : 3b \wedge Sel_{\mathbb{R}}(\pi_0, \pi_1)) \wedge \pi_3 : 3c \wedge Correction(\pi_2, \pi_3) \wedge$
$\pi_7 : (\pi_6 : (\pi_4 : 3d.i \wedge \pi_5 : 3d.ii \wedge Explanation(\pi_4, \pi_5)) \wedge Sel_{\mathbb{R}}(Sel_{\Pi}, \pi_6)) \wedge$
$\pi_{10} : (\pi_8 : 3e.i \wedge \pi_9 : 3e.ii \wedge Confirm\text{-}Question(\pi_8, \pi_9)) \wedge Result(Sel_{\Pi}, \pi_{10}) \wedge$
$\pi_{11} : 3f \wedge Sel_{\mathbb{R}}(Sel_{\Pi}, \pi_{11})$

Here is a gloss of the ULF (4). π_1 is a *complex* discourse unit (CDU) that groups together π_0 and π_1, which are related via some underspecified relation, $\mathscr{R}_1$, which is the target of the correction in π_3. In SDRT, when a correction targets a relation instance, the Correction must take scope over a CDU containing that relation instance. It is also uncontroversial that Explanation holds between π_4 and π_5 and that a Confirmation-Question relation holds between π_8 and π_9. On the other hand it is unclear how to connect the CDU π_6 or the unit π_{11} to the preceding context.[5] The explicit discourse connector *So* signals a result between π_6, π_{10} or some other discourse unit and the CDU π_{10}. However, the left argument of this relation and those of the two underspecified relations are themselves underspecified which we note using the function Sel_Π.

But (4) only represents one possible play in an ME game tree. There could be many more branches. Figure 1 below depicts how (4) and alternative plays branching out from it would form an ME game tree. The relation instances with underspecified arguments are drawn in red.

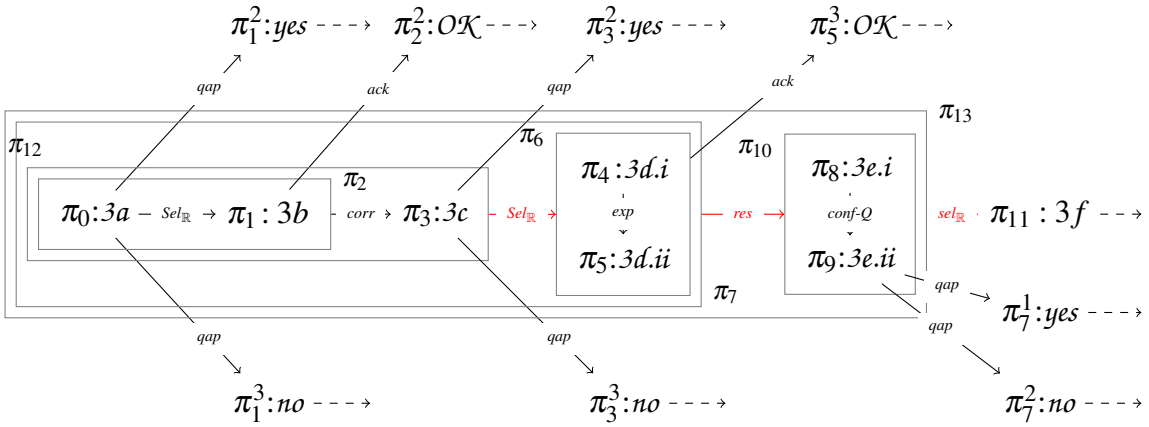

Figure 1: A game tree of plays for (3)

A *history* is a completed SDRS that fills in the underspecified elements of a ULF and thereby serves as an interpretation of a given play. (5) and (6) fill out two histories, $\hbar_1$ and $\hbar_2$, for the observed play ρ and ULF of (4). IQAP stands for Indirect Question-Answer-Pair.

(5) $\pi_{12}: (\pi_2 : (\pi_0 : 3a \wedge \pi_1 : 3b \wedge Background(\pi_0, \pi_1)) \wedge \pi_3 : 3c \wedge Correction(\pi_2, \pi_3)) \wedge$
$\pi_{13}: (\pi_7 : (\pi_6 : (\pi_4 : 3d.i \wedge \pi_5 : 3d.ii \wedge Explanation(\pi_4, \pi_5)) \wedge Correction(\pi_{12}, \pi_6)) \wedge$
$\pi_{10} : (\pi_8 : 3e.i \wedge \pi_9 : 3e.ii \wedge Conf\text{-}Q(\pi_8, \pi_9)) \wedge Res(\pi_7, \pi_{10})) \wedge \pi_{11} : 3f \wedge Backgr(\pi_{13}, \pi_{11})$

(6) $\pi_{12}: (\pi_2 : (\pi_0 : 3a \wedge \pi_1 : 3b \wedge IQAP(\pi_0, \pi_1)) \wedge \pi_3 : 3c \wedge Correction(\pi_2, \pi_3)) \wedge$
$\pi_{13}: (\pi_7 : (\pi_6 : (\pi_4 : 3d.i \wedge \pi_5 : 3d.ii \wedge Explanation(\pi_4, \pi_5)) \wedge Correction(\pi_{12}, \pi_6)) \wedge$
$\pi_{10} : (\pi_8 : 3e.i \wedge \pi_9 : 3e.ii \wedge Conf\text{-}Q(\pi_8, \pi_9)) \wedge Res(\pi_7, \pi_{10})) \wedge \pi_{11} : 3f \wedge Correct(\pi_{13}, \pi_{11})$

To analyze how histories develop from plays, we use the tools of epistemic game theory. We present a detailed, formal development in Asher and Paul (2017), but sketch the essentials

[5] In fact, it is unclear whether the CDU π_6 even exists or whether π_4 alone will relate to the preceding context. We have assumed a CDU for simplicity.

and the linguistic consequences here. In moving from ME games to *epistemic* ME games, we exploit the notion of a *type*, a fundamental tool in epistemic game theory that Harsanyi (1967) used to represent information that players have about each other. In particular, we add to an ME game $\mathscr{G} = ((V_0 \cup V_1)^\infty, \mathscr{J})$, a set of types for the players 0 and 1 and for the Jury $\mathscr{J}$. To model the beliefs of an individual i, which may be a player or the Jury, we also add a function β from a pair of a play ρ, where $\rho \in (V_0 \cup V_1)^\infty$, and type t for i to a probability distribution over types for the other players, types for the Jury, and possible histories (complete SDRSs) given ρ. As the conversation evolves, players will update their beliefs about the history of the conversation and the type of the other players using Bayesian conditionalization over new conversational events they observe (Stalnaker, 2009).

Returning to Example (3), there are two types relevant for interpreting S: the *dishonest* type, t_D, according to which S is trying to cover up the fact that Coleman received the suits but did not declare them, and the *honest* type, t_H, according to which S truly implicates that the Senator did not receive the suits and simply does not want to respond to this charge based on an uncertain source (see (d.ii)). To illustrate how types affect interpretation, we will take the Jury to assign victory conditions in terms of two types: R wins if S's conversational contributions confirm he is of type t_D; S wins if his contributions confirm he is of type t_H. The Jury updates its beliefs about the types of the players as the players make new moves. We now break down the two cases, t_D and t_H, in more detail.

Case 1: Before the start of the press-conference, in the absence of other information, Jury 1's type t is indifferent with respect to S's honesty. That is, the Jury starts with a prior assigning equal probability to t_H and t_D.

Let α be the ULF for (3b). When the Jury updates with the unexpected α as a response to (3a), they are genuinely puzzled by the response. While it's natural to assume that an honest senator has reported every gift he has received, the inference from α to an answer to (3a) (why won't the Senator say who bought the suits?) is complicated and indirect. A Jury must consider the interpretation of (3a) and (3b) conditional on both t_D and t_H. Conditionalizing on α and the assumption that S is of type t_D, the Jury, like R, assigns a high probability to the interpretation illustrated in $\hbar_1$, that (3b) does not answer (3a) and is rather related to it via Background. Conditionalizing on α and the assumption t_H confers only a slightly higher probability to an IQAP relation than a Background relation between (3a) and (3b). When we combine the probabilities over t_D and t_H—because Jury 1 is considering both—we therefore get a higher probability for $\neg$IQAP than for IQAP, leading to higher probability of $\hbar_1$. Conditionalizing in turn on these relative values, the Jury naturally interprets R's response in (3c) as a Correction of S's move in (3b) under the interpretation of (3b) as implicating an answer and therefore satisfying R's request for a direct answer in (3a). In (3d.i), however, S corrects R's Correction, reiterating his original response, and explains why he does so in (3d.ii): the Senator and his staff do not want to comment on unnamed sources on some blog. This would seem to follow whether we conditionalize on t_H *or* t_D.

The upshot of Sheehan's correction should be that (3b) *is* in fact related to (3a) via IQAP. R then picks up on this conclusion and asks a Confirmation Question to confirm that this is indeed

the case. We show this by linking (3e) to the graph built up from (3a)-(3d) with Result in both h_1 and h_2. At this point we could imagine that for our Jury $\mathscr{J}$, $\mathsf{prob}_{\mathscr{J}}(t_H)$ is once again equal to $\mathsf{prob}_{\mathscr{J}}(t_D)$ and h_1 and h_2 are equally likely. But now things go downhill for Sheehan in the eyes of Jury 1. Sheehan in effect refuses to engage with R or confirm the implied result in (3e) by repeating α to every follow up question, Q, on the topic.

Call the exchange in (3a) and (3b) 'round 1' and that in (3c) and (3d) 'round 2'. We now examine how S's responses after round 2 affect the Jury's estimate of his type and its interpretation of what he says. Although S repeats α 10 more times in the press conference from which our excerpt (3) is drawn, for simplicity of this analysis, we shall consider only rounds 3 through 5, for which S has three possible responses to each Q: *yes*, which is short for *Yes, the Senator has received gifts from his friend*; *no* which is short for *No, the Senator has never received gifts from his friend*; and α. The possible continuations for Sheehan that are relevant for the three rounds where *R* repeatedly poses different forms of Q are presented below in tabular form:

	round 3	round 4	round 5
σ_1	*yes*	–	–
σ_2	*no*	–	–
σ_3	α	*yes*	–
σ_4	α	*no*	–
σ_5	α	α	*yes*
σ_6	α	α	*no*
σ_7	α	α	α

Let $\mathbb{S} = \{\sigma_1, \sigma_2, \ldots, \sigma_7\}$. S represents the relevant possible set of plays, over which the Jury has a probability distribution. σ_7 is the actual case in which S responds with α to all instances of Q. An honest senator would have his spokesperson Sheehan respond with a *no* to the reporter's question eventually. To give a precise model for what happens following Asher and Paul (2017), we need some numbers. The belief function for Jury 1 assigns the following probabilities to the sequences of moves σ_n relative to the types t_H and t_D.

	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
t_H	0	0.167	0	0.167	0	0.166	0
t_D	0.125	0	0.125	0	0.125	0	0.125

Notice that at the outset, the Jury accords an equal probability to t_H and t_D; the values on the rows above each sum up to .5.

We can now calculate how conversational moves can affect the Jury's assessment of S's type. In particular, we want to look at two "events", E_H and E_D, representing an honest senator vs. a dishonest senator and defined as follows:

$$E_H = \{t_H\} \times \mathbb{S}, \quad E_D = \{t_D\} \times \mathbb{S}$$

Now let's take a belief function β over the empty play or ULF ε as our starting point to model

the beliefs of the Jury 1, which is of some type t_1, at the start of the press-conference.

$$\beta[\varepsilon](t_1)(E_H) = \beta[\varepsilon](t_1)(E_D) = 0.5$$

The Jury is thus equally unsure as to whether S is honest or dishonest. Recall that the Jury reverts to this distribution after round 2.

Let $\rho_1 = \langle \mathrm{Q} \rangle \langle \alpha \rangle$ be the play after round 2, i.e., after (3d). The strategies that are compatible with ρ_1 are $\mathbb{S}_1 = \{\sigma_3, \sigma_4, \sigma_5, \sigma_6, \sigma_7\}$. Hence, we can define the events

$$E_H(\rho_1) = \{t_H\} \times \mathbb{S}_1, \quad E_D(\rho_1) = \{t_D\} \times \mathbb{S}_1, \quad E(\rho_1) = E_H(\rho_1) \cup E_D(\rho_1)$$

Now,

$$\beta[\varepsilon](t_1)(E(\rho_1)) = 0.708$$

Let $j \in \{4, 6\}$. Then we have, conditionalizing on the new event $E(\rho_1)$:

$$\beta[\rho_1](t_1)(\langle t_H, \sigma_j \rangle) = \beta[\varepsilon](t_1)(\langle t_H, \sigma_j \rangle \mid E(\rho_1)) = 0.167/0.708 = 0.238$$

and for $k \in \{3, 5, 7\}$

$$\beta[\rho_1](t_1)(\langle t_D, \sigma_k \rangle) = \beta[\varepsilon](t_1)(\langle t_D, \sigma_k \rangle \mid E(\rho_1)) = 0.125/0.708 = 0.175$$

Thus after round 3, the belief function of the Jury in Case 1, after Bayesian updates, can be represented in the following tabular form.

	σ_3	σ_4	σ_5	σ_6	σ_7
t_H	0	0.238	0	0.238	0
t_D	0.175	0	0.175	0	0.175

and we have $\beta[\rho_1](t_1)(E_H(\rho_1)) = 0.476$ and $\beta[\rho_1](t_1)(E_D(\rho_1)) = 0.525$

Next, let $\rho_2 = \langle \mathrm{Q} \rangle \langle \alpha \rangle \langle \mathrm{Q} \rangle \langle \alpha \rangle$ be the play after round 3, i.e., after (3f). The strategies that are compatible with ρ_2 are $\mathbb{S}_2 = \{\sigma_5, \sigma_6, \sigma_7\}$. As before, we can define the events

$$E_H(\rho_2) = \{t_H\} \times \mathbb{S}_2, \quad E_D(\rho_2) = \{t_D\} \times \mathbb{S}_2, \quad E(\rho_2) = E_H(\rho_2) \cup E_D(\rho_2)$$

and hence

$$\beta[\varepsilon](t_1)(E(\rho_2)) = 0.587$$

We have, as before

$$\beta[\rho_2](t_1)(\langle t_H, \sigma_6 \rangle) = \beta[\varepsilon](t_1)(\langle t_H, \sigma_6 \rangle \mid E(\rho_2)) = 0.238/0.587 = 0.404$$

and for $j \in \{5, 7\}$

$$\beta[\rho_2](t_1)(\langle t_D, \sigma_k \rangle) = \beta[\varepsilon](t_1)(\langle t_D, \sigma_j \rangle \mid E(\rho_2)) = 0.175/0.587 = 0.298$$

Thus, $\beta[\rho_2](t_1)(E_H(\rho_2)) = 0.404$ and $\beta[\rho_2](t_1)(E_D(\rho_2)) = 0.596$. So after round 2, after Bayesian updates, the Jury believes even more that S has type t_D

The beliefs of the Jury about the type of S (and of the Senator) after each round of the conversation can be represented pictorially as follows.

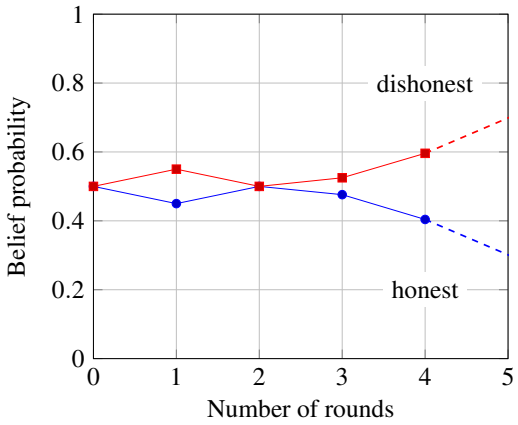

Figure 3

Given these calculations, we can imagine that such a Jury might then stop the conversation once the probability of t_D becomes high enough. For such a Jury, Sheehan's repetitions doom his play to be losing.

A key feature of our analysis is that the Jury's estimation of S's type also affects its *interpretation* of the conversation. The Jury's beliefs about the interpretation of what S says evolve as the Jury conditionalizes on events that are a combination of a new element of play as given in S and an assignment of types. The *yes* and *no* responses in the continuations to (3) would have an unambiguous interpretation as answers to the preceding occurrence of Q. The response α, however, is more problematic; what exactly is the role of α as a response to (3a)? And while α has a natural interpretation in (3d.i) as a Correction, subsequent repetitions of α are even less clear than its use in (3b). Consider, for instance, S's assertion of α as a response to R's question (3e), an affirmative answer to which should be a consequence, were α an indirect answer to (3a). While this and further instances of α could be attached with Background or Correction or with no specified relation at all, the probabilities of these histories given a play are vastly different depending on the type assignments.

To make things more concrete, let's focus just on the two histories we've considered in (5) and (6). Given Jury 1's belief function, repeating α to each variant of Q lowers the probability of t_H. But what does conditionalization do to the probabilities of h_1 and h_2? One plausible hypothesis is that the probability of h_1 covaries with the probability of t_H. Conditionalizing on events of the form $E_D(\rho_i)$ and $E_H(\rho_i)$ thus lowers the probability that repetitions of α after (3d.i) are Correction moves that provide answers to Q and raises the possibility that they are uncooperative moves unconnected with R's questions. An honest Sheehan for this Jury would not have continued to make the α move as a Correction without further explanation. Once S continues to play α, the probability mass shifts more and more to the interpretation of α as a non-cooperative move.

In turn, conditionalization on the event of the interpretation h_1 entrenches the Jury's belief that Sheehan is of type t_D. Thus, histories and types have an important co-dependence. A person's interpretation of a conversation can reinforce or change her beliefs about the players and or the

Jury, which in turn may confirm or change how she shapes the history of the conversation.

Case 2: An alternative Jury, which is strongly predisposed to assign Sheehan t_H would have seen matters differently. Consider the following belief function for this Jury, in which the Jury already has a prior probability of .7 in t_H.

	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
t_H	0	0	0	0.05	0	0.15	0.5
t_D	0.10	0	0.10	0	0.10	0	0

Such a Jury would have already accepted α as a perfectly acceptable indirect answer to (3a) and so opted for the history h_2 for (3). It would also have constructed a different history for the rest of the conversation after (e). It would see each repetition of α as another correction of R's attempts to reopen a topic that Sheehan has already settled. Since S is of type t_H, he need not continue the discussion of a matter that has already been labelled as one that Sheehan will not comment on. Given this interpretation of the repetitions of α, the probability of t_H on the belief function for this Jury would remain high when the Jury conditionalizes upon that interpretation, and S has a winning strategy. See Figure 4 below.

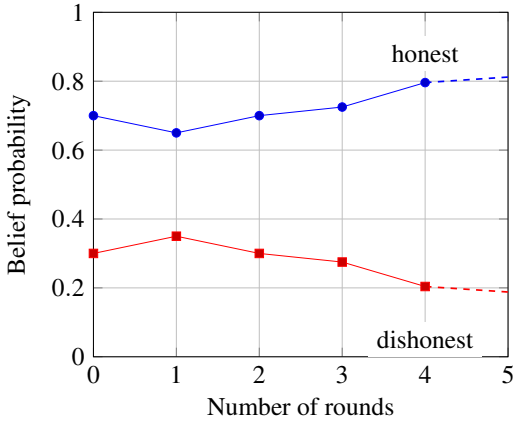

Figure 4

Given the continued high probability of t_H, conditionalizing on events of the form $\{t_H\} \times \mathbb{S}$ in turn assigns a high probability and confirms the continuation of history h_2 in which S continues to correct R. Jury 2 conditionalizes on events like those defined in our analysis of Jury 1 but arrives at a very different conclusion; α is not interpreted as a non-cooperative move but as a Correction, and the updating of probabilities on types and on histories confirm each other.

The co-dependence between an interpreter's assessment of a speaker's type and his interpretation of what she says has several interesting consequences. The first is that *a priori* biases are easy in general to strengthen through interpretation. Furthermore, we have seen that bias is dependent on context. Hence, the model predicts that the more biased talk to which the Jury is exposed, the more it will carry this bias into the interpretation of future conversations. The moral is that interpretation is subject to manipulation. Although we have assumed that few

components of discourse moves are ambiguous, our example illustrates that discourse connections, which are often ambiguous, can lead different Juries to very different conclusions.

This subjectivity goes even beyond language to the interpretation of "facts." Hunter et al. (2017) argue that interpreting non-linguistic events has significant parallels to interpreting conversation and that the two interact to produce a complete picture of a situation. If we couple this view with what we have developed here, our model implies that the interpretation of facts also naturally gives rise to a subjective bias. Figuring out the truth from a collection of facts described to suit someone's purposes is difficult.

The counterpart to this is that being a fair or impartial Jury, being open to other interpretations is not a natural outcome of conversational interaction. The model predicts that constraints exogenous to the natural way speakers interpret conversation and interpretation are needed to lead to an impartial assessment of conversational moves. So how do we ensure fair or unbiased interpretation for the Jury, or anyone else for that matter? This question links our question of assessment of conversational success with the present concern about the subjectivity of interpretation. One important parameter to fix are the types evoked to drive interpretation, as well as the probability distribution over them. In the model of Asher and Paul (2017), the types are simply abstract objects, devices for encoding probabilities about discourse histories and, indirectly, strategies. As such, the set of types is vast and uncountable. Yet typically, only certain types are relevant to interpretation. How do we determine an appropriate restricted set? Doing this also involves bias. To take an extreme case, if we only picked one type in our interpretation of (3), all the probability mass would align on that one type; conditionalization on new evidence would yield nothing new. A Jury considering just the one type would perforce be biased in the extreme and its views impervious to change.

A balanced or fair interpretation would thus need at least two types that are in an intuitive sense contraries of each other, as we have done in our treatment of (3). Moreover, one would need to start out with a balanced distribution over these types or at least be aware of the problems of bias coming from prior interactions. Nevertheless, we strongly suspect that having only two types is in general far from sufficient to arrive at the correct interpretations of conversational moves, and indeed, of the world around us.

4. Analyzing different types of content

A final illustration of the relevance of epistemic games to linguistics comes from the way speakers use *not at-issue* (NAI) information (Potts, 2005). Consider the following dialogue excerpt from the movie *The Princess Bride*, in which Wesley (W) has kidnapped Buttercup (B) and is questioning her about her fiancée Prince Humperdink (H). Buttercup and Wesley were once in love, but Wesley has disguised himself as a pirate, so Buttercup does not recognize him. They have just noticed their pursuers: Humperdink and his men.

(7) a. **B**: He (Humperdink) can find a falcon on a cloudy day, he can find you!
b. **W**: So you think your dearest love will save you?
c. **B**: I never said he was my dearest love and yes, he will save me. That I know.

d. **W**: You admit to me that you do not love your financé?
e. **B**: He knows I do not love him.

As a background to (7), it is clear from the context that Wesley's goal is to determine Buttercup's feelings for Humperdink while concealing his identity. Concealing his identity is important because Buttercup would have good reason to hide her love for Humperdink from Wesley.

Wesley uses a presupposition in (7b), a form that conveys NAI content, to try to discern whether Buttercup loves Humperdink. As many linguists have noted, NAI constructions do not admit of simple rejoinders; Buttercup cannot take issue with the presupposition that Humperdink is her dearest love with a simple 'no', 'I disagree', or 'that's not true'. Getting Buttercup to reveal her type in this situation involves a more complicated strategy than simply asking directly if she loves Humperdink. Why does Wesley choose such a strategy? Why does he pretend to play along with Buttercup's apparent conversational goal—to convince Wesley that he should let her go because he is going to be caught—when he isn't actually interested in it?

We believe epistemic ME games can help clarify matters, but first, we need some background assumptions. We follow Hunter and Asher (2016)'s discursive, SDRT analysis of NAI content according to which NAI content results from the way that an utterance contributes to a hierarchical discourse structure. At any given point in a discourse, certain discourse moves will be more salient and easier to build off of than others. The set of salient discourse units, referred to as *The Right Frontier* in various discourse theories, will evolve as a discourse proceeds in a way that is subject to general discursive principles. Hunter and Asher's central claim is that utterances that involve multiple discourse units, which they argue must be the case when an utterance contains both NAI and AI content, have their own internal discursive structure, and the same general discursive principles that determine whether other discourse units are on the Right Frontier or not also determine which parts of a multi-part utterance are on the Right Frontier and therefore salient. Speakers exploit these principles to make certain parts of their utterances easier to build off of than others. Those parts that are on the Right Frontier will be at-issue (so long as they are on the RF); those that are not will be NAI. In SDRT, presupposed content is attached to a position off the Right Frontier (Asher and Lascarides, 1998); the presupposition of (7b) is therefore NAI, as desired. With Buttercup's correction in (7c) and Wesley's subsequent follow-up in (7d), however, the discourse shifts seamlessly to one in which the presupposed content is placed back on the Right Frontier and thereby made at-issue.

With this background in place, let's look more closely at the discourse structure of (7) and its interaction with conversational goals. Suppose that Buttercup is of one of two types: either she loves H and is of type t_l, or she does not and is of type t_n. We will suppose that in either case, she is not interested in dissembling her type. Wesley's conversational goal can be formulated as getting Buttercup to reveal whether she is of type t_n or type t_l while concealing his type. We assume further that after (7a), Wesley has the option of using (7b) with the NAI device or asking the direct question DQ: *Do you love Humperdink?*. In the context of (7a) and the more general pursuit by Humperdink and his men, DQ would blatantly shift the topic and naturally arouse suspicions: why would an unknown pirate care about Buttercup's affections for Humperdink? There is a high probability that Buttercup would fashion a conversational continuation—e.g.,

why should YOU care about that?—that would endanger Wesley's goal of concealing his identity. By opting for an NAI construction, Wesley makes *Buttercup* responsible for the shift in discourse topic, which is less likely to arouse suspicion.

The expected utility of DQ is thus lower than that of (7b) when it comes to achieving Wesley's goal of concealing his identity. This goal in turn serves the larger goal of determining Buttercup's type. Buttercup would have no obvious reason to hide her love for Humperdink from a total stranger, especially a stranger who knows that she is engaged to Humperdink. On the other hand, her old feelings for Wesley, whom she believes at this point to be dead, might give her reason to be less than forthcoming with him. Buttercup's rejection of the NAI content in (7b) is more significant, and therefore more useful for determining Buttercup's type, if she does not realize she is talking to Wesley. Thus the NAI device in (7b) has a higher expected utility for Wesley's primary goal as well.

Now suppose that Buttercup had been of type t_l and had not questioned the NAI content, responding only to the AI content with something like, *Yes, he will save me* (Sv). By conditionalizing on Sv, which entails Buttercup's acceptance and public commitment to the NAI content[6], the probability that Buttercup is of type t_l is plausibly high. The NAI construction would be helpful in this case as well, because it would allow Wesley to cut his losses and infer that Buttercup's romantic allegiances had shifted, another way of obtaining his winning condition.

We have illustrated one example of a strategic use of NAI content, but there are many others (see Hunter and Asher, 2016). In each of these cases, NAI constructions are strategically useful when a speaker s_0 has a conversational goal that he wishes to conceal. Our model predicts this strategic use of NAI devices if we make the reasonable assumption that a move by s_0 to place a bit of content ϕ on the Right Frontier as a topic for discussion raises the probability that s_0 has a conversational goal of getting a commitment from his interlocutor, s_1, to ϕ or $\neg\phi$. Someone wanting to conceal his desire to extract such a commitment will be better off placing ϕ in an NAI position. In this case, he can still get the commitment he seeks: if s_1 continues the conversation without disputing ϕ, s_1 commits to the content of ϕ. Conversely, if s_1 does dispute ϕ,[7] then she may reveal information to s_0 without s_0 making ϕ discourse salient.

We formalize the strategic use of NAI content in Proposition 1. Recall that $\Diamond\phi$ is true on a conversational string just in case ϕ is true at some stage in the conversation; let $C_1\phi$ stand for 's_1 commits to ϕ,' and let '$\models$' stand for the satisfaction relation defined over the basic exogenous semantics for plays in $(V_0 \cup V_1)^\infty$ (for details, see Asher et al., 2016). Let ρ_n for $\rho \in (V_0 \cup V_1)^\infty$ be a prefix of length n of ρ and let β_1 be the belief function for a fixed type of s_1. Let t_ϕ be a type of s_0 whose winning condition includes a move that commits s_1 to ϕ. Let ε be the empty play in $(V_0 \cup V_1)^\infty$.

Proposition 1. *Let G be an ME game with* $Win_0 \subseteq \{\rho \in (V_0 \cup V_1)^\infty : \rho \models \Diamond C_1\phi\} \cap \{\rho \in (V_0 \cup V_1)^\infty : \neg\exists n \beta_1[\rho_n](t_\phi) > \beta_1[\varepsilon](t_\phi)\}$. *Then if* 0 *has a winning strategy* σ *in G,* σ *will include a use of an NAI device for conveying a question about* ϕ *instead of an AI device for conveying* ϕ.

[6]Wesley's use of a presupposition should not entail that Buttercup commits to its content, but once she opts to build on Wesley's structured discourse contribution, she commits herself to the whole of both its content and its structure. See Hunter and Asher (2016) for more details.

[7]For an account of corrections, see Asher and Lascarides (2003).

Hunter and Asher (2016) argued that the NAI status of a bit of content ϕ cannot be attributed to syntactic and semantic features of ϕ alone; its status is ultimately determined by how ϕ attaches to a larger discourse structure. Our discussion here makes a similar point about conversational goals: the nature of a goal cannot be recovered directly by considering the at-issue status of content in the discourse; its nature is ultimately determined by the larger discourse structure or history. In particular, we cannot assume that the difference between AI and NAI content is that AI content directly addresses a conversational goal while NAI content plays some secondary, supporting role relative to conversational goals, contra claims made in by Roberts (2012) and Simons et al. (2010) *inter alia*. If we consider only (7a)-(7c), then it is true that Wesley's apparent conversational goal aligns with the AI content of his utterance in (7b); that is, it seems that he shares Buttercup's goal to determine whether Humperdink will catch him. However, once we consider the longer string including not only (7d) and (7e), but also Wesley's subsequent questioning of Buttercup, then we understand that his actual goal is more accurately reflected by his NAI moves. Even if Buttercup is not immediately aware of this actual conversational goal, the audience of the film, acting as a Jury, can see this clearly.

5. Conclusions and future work

We have elaborated some consequences of using ME games for linguistic analysis, consequences that we think touch on important issues in pragmatics and semantics. We have shown how a notion of conversational success and a precise definition of conversational goals can affect the discourse structure of a dialogue. We have also shown how to add epistemic considerations to discourse interpretation, which has allowed us to formulate an analysis of the power and the limits of subjectivity in interpretation. We have further shown how the framework of epistemic ME games yields an analysis of bias, its ubiquity and its effects on interpretation. At the same time, we have argued that there are limits to bias; one cannot arbitrarily reinterpret unambiguous messages to mean something different from what they normally mean.

Another very important issue that we leave for future work concerns the notion of a fair or "unbiased" interpretation, something of interest not only to linguists but the general public at large. While such a notion has an intuitive meaning, working out a precise analysis within a framework like epistemic ME games requires an analysis of the types relevant to such an interpretation. We also feel that fair interpretations are hard. Given that a fair interpretation is a conversational goal, our framework can actually tell us precisely how hard it is, but we do not know at present the complexity of such a goal nor even how to formulate it precisely.

Another application of epistemic ME games that we find linguistically interesting is their use in analyzing the strategic uses of vehicles of content that affect discourse salience or that introduce content without affecting salience. To this end, we have given a preliminary analysis of the strategic usefulness of not at-issue expressions of content. Future work will aim to improve this analysis with more case studies and a more careful taxonomy of expressions that affect discourse salience.

References

Asher, N. and A. Lascarides (1998). The semantics and pragmatics of presupposition. *Journal of Semantics 15*(2), 239–299.

Asher, N. and A. Lascarides (2003). *Logics of Conversation*. Cambridge University Press.

Asher, N. and A. Lascarides (2013). Strategic conversation. *Semantics and Pragmatics 6*(2), 1–62.

Asher, N. and S. Paul (2016, July). Evaluating conversational success: Weighted message exchange games. In J. Hunter, M. Simons, and M. Stone (Eds.), *20th workshop on the semantics and pragmatics of dialogue (SEMDIAL)*, New Jersey, USA.

Asher, N. and S. Paul (2017). Epistemic message exchange games. In progress.

Asher, N., S. Paul, and A. Venant (2016). Message exchange games in strategic conversations. *Journal of Philosophical Logic*. doi.org/10.1007/s10992-016-9402-1.

Crawford, V. and J. Sobel (1982). Strategic information transmission. *Econometrica 50*(6), 1431–1451.

Franke, M. (2008). Meaning and inference in case of conflict. In K. Balogh (Ed.), *Proceedings of the 13th ESSLLI Student Session*, pp. 65–74.

Franke, M., T. De Jager, and R. Van Rooij (2012). Relevance in cooperation and conflict. *Journal of Logic and Computation 22*(1), 23–54.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and Semantics Volume 3: Speech Acts*, pp. 41–58. Academic Press.

Harsanyi, J. C. (1967). Games with incomplete information played by bayesian players, parts i-iii. *Management science 14*, 159–182.

Hunter, J. and N. Asher (2016). Shapes of conversation and at issue content. In *Semantics and Linguistic Theory 26*, pp. 1022–1042.

Hunter, J., N. Asher, and A. Lascarides (2017). A formal semantics for situated conversation. unpublished manuscript.

Lamport, L. (1980). Sometime is sometimes not never: On the temporal logic of programs. In *Proceedings of the 7th ACM SIGPLAN-SIGACT symposium on Principles of programming languages*, pp. 174–185. ACM.

Lewis, D. (1969). *Convention: A Philosophical Study*. Harvard University Press.

Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.

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

Simons, M., J. Tonhauser, D. Beaver, and C. Roberts (2010). What projects and why? In N. Li and D. Lutz (Eds.), *Semantics and Linguistics Theory (SALT) 20*, pp. 309–327.

Spence, A. M. (1973). Job market signaling. *Journal of Economics 87*(3), 355374.

Stalnaker, R. (2009). Iterated belief revision. *Erkenntnis 70*(2), 189–209.

van Rooij, R. (2004). Signalling games select horn strategies. *Linguistics and Philosophy 27*, 493–527.

Venant, A. and N. Asher (2015). Ok or not ok? In *Semantics and Linguistic Theory 25*. Cornell University Press.