Grice, Utterance Choice, and Rationality*

Daniel Rothschild

1 The Status of Gricean Explanations

This paper is about pragmatic explanations of certain linguistic phenomena based on Grice's theory of conversational implicature. According to Grice's theory, audiences draw inferences about what the speakers are communicating by reference to the assumption that speakers follow certain *maxims* governing cooperative conversation. Such explanations of the inferences audiences make about speakers play a central role in many parts of philosophy of language and linguistics.

Pragmatic explanations in terms of conversational implicatures are usually preferred to other sorts of explanations. This practice is justified by one or both of these two ideas: a) that Grice's theory of conversational implicature is a well-confirmed empirical account of aspects of language use, or b) that Grice's theory of conversational implicature actually *follows* from more basic assumptions about the speech-act situation (such as the rationality of speakers and their goals). If either of a) and b) is right then Gricean explanations, if they are sufficiently predictive, should be sought where possible.

*Many thanks to Chris Barker, Eliza Block, Jessica Boyd, Richard Breheny, Arudra Burra, Nathan Klinedinst, Michael Rothschild, and Jacob Rosen for help with various bits of this. I'm particularly grateful to Michael Franke and other members of the "Philosophy of Language and Beyond" reading group at the ILLC for their general comments as well as for pointing out two significant mistakes in an earlier version. Many thanks, also, to an audience at the Centre for Metaphysics and Mind at the University of Leeds.

This paper is not about a). As for b): surprisingly there's been rather little detailed discussion of the degree to which Grice's theory of conversational implicature follows from plausible assumptions about communication. Certainly many people have thought and said things to the effect that Grice's maxim of quantity ("make your contribution as informative as possible") simply follows from the fact that speakers are rational agents with the goal of communicating as much information to the audience as possible. Nonetheless, very few people have explained exactly what this claim amounts to or tried to justify it with much rigor. Another way of seeing whether Gricean explanations are stipulative or not is by working out whether the maxims are necessary at all in the derivation of particular implicatures, or whether implicatures can be worked out by way of more basic assumptions instead. Again there are not many sufficiently precise discussions of this question.[1]

This paper focuses on *scalar implicatures* which have received a lot of attention in the pragmatics literature over the years. The plan is as follows: First, I isolate a means of deriving certain scalar implicatures by way of a version of the maxim of quantity (and a few other plausible assumptions). Then I ask the question whether the maxim of quantity was really necessary in the derivation or whether the same derivation could have been accomplished instead simply by reference to assumptions about the rationality and the goals of the participants in speech-act situations. My conclusion is that in some simple cases the fact that speakers apply the Gricean maxims is predicted by basic assumptions of rationality and the desire to be informative, but that in more complex

[1]Notable exceptions occur in the game theory and pragmatics literature (e.g. Benz et al., 2005). The basic problem, for our purposes, with much of the game theory literature is that the goal of the enterprise tends to be to show how to derive Gricean implicatures using a game-theoretic approach, not to *assess* the extent to which the implicatures are rationally compelling. Although I will use game-theoretic representations here, this paper is not an attempt at *grounding* Gricean reasoning using concepts borrowed from economics, rather it's an attempt at using those concept to assess the degree to which Gricean reasoning is rationally compelling. The papers in the game theory literature perhaps most closely related to my concerns are de Jager and van Rooij (2007) and van Rooij (to appear), however for various reasons I'll touch on later I do not find even that these papers do much to answer the exact questions I'm posing.

cases rationality does not predict the application of the Gricean maxims.

1.1 Scalar Implicatures using the Maxims of Quantity and Quality

Sentence (1) triggers a simple instance of a scalar implicature.

(1) John passed some of the students.

An utterance of (1) would naturally lead the audience to believe that the speaker did not believe (2) (not that he believes it is false, just that he doesn't believe it is true):

(2) John passed all of the students.

There are probably some assumptions required for the audience to really be entitled to draw this conclusion. For instance, that the truth of (2) rather than (1) is relevant, that the speaker is trying to be helpful and wants the audience to know what he does. But if these assumptions are in place, it is natural for the audience to assume that a speaker who utters (1) does not believe (2). (It is called a *scalar* implicature because it involves the relation between terms like "some" and "all" that are often said to be part of a scale of strength (Horn, 1972).)

The Gricean story about this inference goes as follows: The speaker said (1). He could have said (2) instead. (2) entails (1). If he had said (2) it would have been more informative. The maxim of quantity (limited by constraints on quality (("say only what you believe to be true")) and relevance) says that we should utter the most informative, relevant statement that we believe to be true. So if the speaker had believed (2) he would have said it. Therefore he does not believe (2).[2]

[2] Or, anyway, it goes *something* like that.... Often in the course of this paper, I'll assume without mentioning that various propositions are relevant to both the and the speaker and audience as well as

A well-known problem with this sort of argument is that this general inference scheme is too strong.[3] For instance, it is equally true the speaker could also have said:

(3) John passed some but not all of the students.

However, an utterance of (1) most definitely does not lead to the inference, generally, that the speaker does not know that (3) is true—in fact an utterance of (1) often leads to the inference that the speaker *does believe* that (3) is true.

Some have tried to respond to this by appealing to the maxim of manner, which enjoins the speaker to "be brief", however, it is not clear how exactly this works. Certainly there is no way, based on these simple formulations of the maxims alone, to *derive* the implicature associated with (2) but not the one associated with (3). Maybe there is some way of spelling out the way the maxims interact that will give the right results, but someone will actually have to do this to provide a Gricean story of the inference from an utterance of (1) to the fact that the speaker does not believe (2). Such a story would amount to an algorithm for generating quantity implicatures. I don't know of any such algorithm, nor am I convinced that such an algorithm is possible (at least possible to formulate in such a way that it is even plausibly grounded in the various Gricean maxims). Instead, in the next section, I'll discuss a way of side-stepping this issue entirely.[4]

1.2 Scalar Implicatures as Lexical Choice

Speaking a language is a bit like riding a bike. When you ride a bike, certain decisions you make present themselves as *choices*. In particular, if you are paying attention, you

the fact that the speaker actually *wants* to convey as much information as possible to the hearer, since it gets a bit exhausting to continually reiterate all these assumptions.

[3]The problem is discussed by Fox (2006), it is attributed to class notes of Irene Heim and Kai von Fintel and Kroch (1972).

[4]I am indebted to Eliza Block for many discussions of this issue.

choose whether to shift up or shift down the gears of the bike. Other equally important parts of bike riding, such as how much you lean into turn, do not present themselves as choices even though what you do is under conscious control. In the activity of speaking there may also be certain choices you make: For instance whether you say "walked" or "ran", in a sentence of the form "John ___ to the post office" is perhaps sometimes a salient choice for you. Likewise, whether you should say "some" or "all" in a sentence of the form "John passed ___ of the students" may also be a choice that presents itself to you. Perhaps, however, whether to say "some" or "some but not all" in that sentence is *not* a choice that generally presents itself to you.

In other words, when you expresses a thought in speech, there are a myriad of different ways you could do so, but only some of the decisions come in the form of psychologically prominent choices. Now, one could suppose (and to do so is an empirical supposition) that some choices salient to the speakers are guided by the Gricean maxims. In particular, we could suppose that when one says a sentence such as (1) it is a salient choice to say "some of the students" rather than "all of the students", and that this choice is guided by the Gricean maxims. It may be that Gricean maxims applied *globally* are simply too vague to yield any conclusions about what speakers mean when they speak, but that because of the way language use is structured, we often apply them to local, salient choices.

Following Horn (1972), I will assume that at least in some instances the choice of which of a closed class of scalar items to use is itself subject to the Gricean maxims (independently of all the other choices made in uttering a sentence). In this case, with some plausible assumptions, we can derive the implicatures we want. When a speaker utters (1) he is choosing to utter "some" rather than "all", hence he is choosing to utter (1) rather than (2), and this limited choice is governed by Grice's maxims. The maxim of quantity enjoins him to say as much relevant information as possible. Let us assume that this amounts to a directive to utter the sentence with the strongest truth-conditions that you believe (all else equal). So it follows that if the speaker had believed (2) he would have said it. It follows that since he said (1), he does not believe (2) (i.e. he lacks belief in the truth (2), he may also have believed it is false). Thus the audience can infer that a speaker who says (1) does not believe (2). (This inference is what we're calling the implicature—though Grice probably wouldn't have called such an unadorned inference an implicature.)

2 Does the Maxim of Quantity Follow from Other Assumptions?

I am not sure whether or not this manner of deriving scalar implicatures is correct or not. It is quite similar to the best contemporary accounts (Spector, 2006; Schulz and van Rooij, 2006; Sauerland, 2004). However, it is not my intention to argue that this way of doing things is right. Rather I want to ask this question: supposing Gricean reasoning does operate on certain lexical choices, then do the Gricean maxims (in particular the maxim of quantity), as applied in these particular situations, follow from more basic assumptions about the communicative situation? To be concrete consider the above explanation of the scalar implicature from an utterance (1) to the fact that the speaker does not believe (2). There we assumed that the speaker was obeying the maxim of quantity in the sense that he was constrained to utter the most informative sentence that he could (consistent with his beliefs). Is this assumption, in this particular situation, eliminable? Could we instead give an explanation that only made reference to some more basic facts, such as the goals and rationality of the speaker and audience?

The answer to this question might seem very obvious. Here is a way of trying to reproduce the Gricean reasoning about an utterance of (1) without explicitly using the maxim of quantity: Both the speakers and the audience are rational agents. Moreover, in the cooperative conversational situation the speaker has relevant information that the audience lacks. It is one of the speaker's goals to convey as much relevant information as possible to the audience. Now if we assume that the truth of (2) rather than (1) is relevant it would seem obvious that a speaker who has the choice of saying (1) or (2) should say (2) (all else equal).

As it happens, in this particular case this conclusion, and indeed this *sort* of line of reasoning, is correct. However, there is a clear respect in which this sort of reasoning is deficient in general. The problem is this: A rational agent wants to *convey* as much information as possible. However, we cannot assume that he wants to say the sentence whose literal meaning is the strongest. For the rational agent should not care whether he conveys his information by direct assertion or by implicature. So the fact that the speaker wants to convey as much information as possible does not immediately yield the conclusion that the speaker will always say the truth-conditionally strongest sentence in a choice situation. In other words, if we formulate the maxim of quantity as a constraint on which sentence a speaker should utter (i.e. the most informative one the speaker believes) then it does not automatically follow from the simple goal of trying to convey as much information as possible.[5]

To derive the maxim of quantity (as it applies in this instance) we need to take for granted the truth-conditional conventions of the language (i.e.the semantics of the language and the convention of assertion), and show that a speaker cannot convey (by any means) more relevant information that he believes by saying (1) rather than (2). This will be a bit laborious, but it is worth the effort since we need to see why these

[5]Note that Grice's own formulation is ambiguous: "make your contribution as informative as is required for the purposes of the conversation". Does your contribution mean what is said by your utterance or what is said *and* what is implicated? When I refer to the maxim of quantity I mean "make your truth-conditional (pre-implicature) contribution as strong as possible". If you formulate it the other way, then you have to phrase the dialectic of this paper differently, but nothing important changes.

explanations *succeed* in simple cases, in order that we can later see why they *fail* in more complex cases.

In order to make our explanation work we need to assume that there are not extra implicatures (not based on quantity-type reasoning) arising from (1) but not (2) which would give the speaker an independent reason for saying (1) rather than (2). Let us divide the possible states of mind of an utterer of (1) or (2) into two types: those who believe (1) but not (2) and those who believe both (call them type 1 and type 2, respectively). As a way of limiting the possible implicatures (and thus making our problem solvable) we will assume that the only possible information being directly conveyed is about whether the speaker is of type 1 or type 2 (by "directly" I don't mean conveyed by assertion rather than implicature, I mean conveyed first rather than as the logical consequence of something else that is conveyed).[6] I will also assume that the audience assigns a positive subjective probability to the proposition that the speaker is of type 1 and to the proposition that the speaker is of type 2 (i.e. he has not ruled out either option in advance of the speech act).

Now that we have laboriously spelled out the situation, it is easy to see why the Gricean explanation really does come for free. The conventions of the language and assertion (i.e. say only what you believe to be true) allow a speaker of type 2 to say either (1) or (2), but they allow a speaker of type only 1 to say (1). A rational audience when he hears an utterance of (1) will have to put some degree of credence in the proposition that the speaker is of type 1 (since we must assume he has not ruled that possibility out to begin with). A rational audience who hears an utterance of (2) will also have to believe with full credence that the speaker is of type 2 (since a speaker of

[6] This is not an unreasonable assumption: lots of other information can be conveyed by utterances of (1) and (2), but it typically follows from the type of the speaker plus other assumptions. For instance, that (1) is true follows from the speaker being of type 1 and his beliefs being true. That (2) is false follows from the speaker being of type 1 and him not being ignorant about the truth of (1) or (2). See Spector (2006), Schulz and van Rooij (2006), and Sauerland (2004) for extensive discussion of these latter inferences based on speaker *expertise*.

type 1 cannot say (2), given the semantics of the language and the norms governing assertion). It follows that any rational speaker of type 2 who wants to convey as much information to the audience as possible will need to utter (2) rather than (1) given the choice.

2.1 Game-Theoretic Interlude

The situation described above can be modeled more explicitly using the resources of game theory. Game theory has the theoretical apparatus to describe decisions that involve more than one agent: conversational situations are very naturally thought of as an interaction between two players, the speaker and the hearer.[7] Moreover, game theory has *equilibrium concepts* which give a precise characterization of pairs of strategies (ways of acting in the game for both players) that have certain desirable features (such as being mutually rationalizable in certain respects).

It is best to think of speech games as a special kind of two-player games called *signaling games*. A signaling game has essentially three stages: first the sender gets sent a message (by nature) revealing his type, then the sender performs an action (he sends a message), and then the receiver performs an action. The crucial feature of the signaling game is the *informational asymmetry*. The sender but not the receiver knows what type he is, so the hearer when he responds to the sender only knows what action is performed by the sender not (directly) what type the sender is. A second feature, special to our game, is that it is a cheap-talk game. This means the payoff to both players depends just on the type of the sender and the action of the receiver, but not at all on the action of the sender (i.e. the message the sender sent). (Some terminology: each move that is made by any player—including nature—puts us is in a new *node*, and corresponding to every set of moves made up to some point in the game there is separate node.) A

[7] It is the way Lewis (1969) thought of them.

strategy for either player is a complete specification of the action the player performs at each node of the game (although in any given game, some nodes will not be reached). (More terminology: probabilistically specified strategies are called *mixed*, deterministic ones are called *pure*.)

All this is a bit abstract, so let us move to the specifics of the game we are interested in, which is actually a very special variant on the basic form of the signaling game (in some sense it is in fact not a signaling game at all, though it is closely related to one). We will assume a couple of other things. In our game nature assigns the sender (henceforth, the speaker) either type 1 or type 2 (which are the same as in the previous section). The receiver (henceforth, the audience) has antecedent beliefs about the probability that the the speaker is of type 1 or type 2 and the speaker knows these and neither is zero. There are two actions available to speakers of type 2: utter (1) or (2), and one action available to speakers of type 1: utter (1). A special feature of our game is that the receiver does not perform an action, per se, but rather the payoff to both players simply depends on the degree of credence the audience puts in what type the speaker is. We won't need to get too specific about the payoff function, but we'll assume that it is ordered by the degree of credence the audience has in the true proposition about the type of the speaker. In other words, both players get the same payoff and it depends just on how much credence the receiver puts in the truth about the sender (the more, the better). Figure 1 diagrams the options for speakers of different type. Obviously, there's no choice of action for the audience, just changing belief states about the speaker.

The standard solution concept for signalling games (which this is a variation on) is that of a perfect Bayesian equilibrium.[8] Given the somewhat different character of the game I am considering from the standard signaling game, I need to substantially adapt

[8]Gibbons (1992) is a quick and clear introduction to signaling games and perfect Bayesian equilibria; thicker books on game theory go into more gory detail.

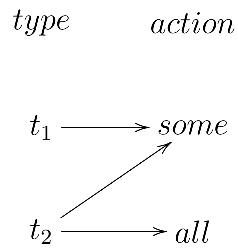

Figure 1: Simple "some"/"all" game

the equilibrium concept.[9]

Variant of Perfect Bayesian Equilibrium (PBE) A speaker strategy s is part of a *perfect Bayesian equilibrium* of the speaker strategy and audience's belief dynamics if the following holds:

1. The speaker knows what type he is prior to his action and the audience has credences at all points in the game as to what type the speaker is.
2. The audience updates his beliefs as the game progresses according to Bayes' rule and his belief that the speaker is playing s (insofar as the latter is possible to maintain).
3. The speaker's action after he has been assigned a type is rational in light of his type and the fact that the audience will update his beliefs according to rule 2.[10]

Let me make some comments. First of all, condition 3 ensures the Nash-style to the equilibrium: the speaker has no incentive to switch his strategy. Given that the payoffs

[9] I hope, however, it will be recognizable as pretty much the only way to adapt the perfect Bayesian equilibrium concept to this sort of game. Michael Franke has stressed to me the difference between this type of game and the standard signaling games, and hence the difference between this type of equilibrium strategy and the normal prefect Bayesian equilibrium concept.

[10] Note that given a speakers strategy, how the audience updates his belief simply *follows* from his antecedent belief in the probability of the different speaker types and s, so there is no real "choice" of audience action. Nonetheless, there is a real sense in which the audience's belief updating and speaker choice are in equilibrium since the former must be rational in light of the latter *and* the payoffs.

are based on the correctness of the audience's belief, the updating of beliefs based on Bayes' rule will also be the best "response" to the speaker's strategy. Second, note that there are no tricky assumptions of common knowledge or anything: to check if s is a PBE all you need to do is calculate what the audience's credences will be after each message (using Bayes's rule) and then check that s is rational for each type of speaker in light of those credences (and the payoffs associated with them).

More generally, note what this solution concept is getting at: If a strategy is a PBE (in this variant) then it is a speaker strategy such that if the rational audience believes the speaker is playing the strategy, then the speaker has no incentive to deviate from it. This is, in a non-technical sense, makes the strategy *rationalizable*.

It is easily provable that only one sender's strategy will yield a PBE: the strategy where the sender of type 2 utters (2) and the sender of type 1 utters (1).

Proof. Suppose a speaker's strategy s is such that there is a positive probability that the speaker of type 2 utters (1). If the audience knows the speaker's strategy is s, by Bayes's rule, the audience must give nonzero credence to the proposition that the speaker is of type 1 when the speaker utters (1), given that the audience had some credence before that the speaker might be of type 1. In this case the speaker is better off altering s so that he always utters (2) when he is of type 2 since the audience will always respond to an utterance of (2) by assigning full credence to the proposition that the speaker is of type 2.[11] So s is not a PBE. □

In our game, then, there is only one PBE. Does that mean that it is *rationally compelling* that the speaker use that strategy when making a choice between (1) and (2). I take it that might be a *defeasible* assumption that the unique PBE is really the

[11]Note that I assume that when the audience receives a message that isn't expected on s he updates his belief in a way compatible with Bayes' rule and the structure of the game. Otherwise if s was the pure strategy of always uttering (1) then it would not be determinate how the audience would react.

only rational thing to do. In this case we don't need to rely on the equilibrium concept to see why people should act as they do. For the reasons explained it's plainly irrational for speakers of type 2 to do anything but utter (1), and that is sufficient to show that they will act according to the Gricean strategy.[12]

3 Slightly More Complex: 3-Point Scales

We just considered a choice between two utterances (deriving from a lexical choice between "some" and "all"). Now consider a choice between three sentences (deriving from a lexical choice between "some", "most", and "all"):

(4) John passed some of the students.

(5) John passed most of the students.

(6) John passed all of the students.

Corresponding to these three sentences are three types of speaker: type 1 believes only (4), type 2 believes only (4) and (5), type 3 believes all three statements. Informally there are pretty good arguments why speakers of type 3 should say (6): However the audience interprets an utterance of (4) or (5), it seems likely that the audience will put some credence into the proposition that someone who says (4) or (5) is *not* of type 3. Thus it would seem that a speaker of type 3 should utter (6), since that is *guaranteed* to make the audience have the right belief about him. If speakers of type 3 always utter (6) then the rest of the choice looks like the previous game and so we should expect speakers of type 2 to utter (5). This anyway is the quick, informal version. But as we

[12]It might be useful to compare the discussion of van Rooij (to appear) to this section, he considers a related game and comes up with a similar conclusion, using some different equilibrium concepts more appropriate to his particular game. Our approaches differ more substantially for the remainder of the paper.

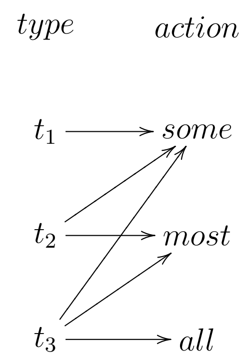

Figure 2: "some"/"most"/"all" game

shall see formulation this reasoning in a rigorous and compelling manner is not that easy.

Let us look at the utterance choice between (4)–(6) as a game again. The possibilities of utterance choice are shown in Figure 2. Again the utterances are limited by the semantics of the language and the convention that assertion is restricted by belief. Again we assume that the audience forms a belief about the speaker's type after the speaker acts, and we assume the payoffs (the same for both players) are ordered by the audience's degree of credence in the truth about the speaker. (As before, we're treating the maxim of quality as a convention governing assertion.) It is easy to show that there is more than one perfect Bayesian equilibrium for this game. I will describe two pure speaker's strategies that make different equilibria (Figures 3 and 4).

Gricean Speakers of type 1 say (4), speakers of type 2 say (5), and speakers of type 3 say (6).

Partial Pooling Speaker of type 1 say (4), speakers of type 2 say (4), and speakers of type 3 say (5).

That the Gricean strategy is part of a perfect Bayesian equilibrium is obvious. But why is the Partial Pooling strategy? To see this we need to think about what deviations

$$\begin{tikzcd} \textit{type} & \textit{action} \\ t_1 \arrow[r] & \textit{some} \\ t_2 \arrow[r] & \textit{most} \\ t_3 \arrow[r] & \textit{all} \end{tikzcd}$$

Figure 3: The Gricean strategy for“some”/“most”/“all” game

$$\begin{tikzcd} \textit{type} & \textit{action} \\ t_1 \arrow[r] & \textit{some} \\ t_2 \arrow[ur] & \textit{most} \\ t_3 \arrow[ur] & \textit{all} \end{tikzcd}$$

Figure 4: Partial Pooling strategy for“some”/“most”/“all” game

the speaker could make from the pooling strategy and demonstrate that they will not improve his payoffs assuming the audience believes the speaker is playing the pooling strategy. Speakers of type 1 have no choice. Speakers of type 2, on the pooling strategy, say (4) but they could instead say (5). However, saying (5) will reduce their payoffs since the audience will then believe they are of type 3. Speakers of type 3 already get the audience to believe they are of type 3 with certainty, on the pooling strategy, so they have no incentive to change their strategy (of course they could just as well say (6) so it is not a *strict* perfect Bayesian equilibrium, as the Gricean strategy is).

3.1 Why the Gricean Strategy is Compelling

Even though there are two perfect Bayesian equilibria, one of them may still be rationally compelling to speakers and audiences in this particular situation. We can first note that the Gricean equilibrium has certain advantages over the Partial Pooling equilibrium. Importantly, players are better off (or at least no worse off) playing the Gricean strategy, no matter what the type of the speaker. Also, the Gricean strategy is a strict equilibrium, but the Partial Pooling strategy is not. (In other words, deviating from the Gricean strategy always makes the players worse off when the other player plays according to the Gricean strategy). However, it is not clear that either of these facts makes the Gricean strategy rationally compelling in any way.

I think, however, that in the speech-act situation we are considering the Gricean strategy *is* rationally compelling. In particular, I will argue that given that speakers have not, in any sense, agreed ahead of time to play according to any particular strategy and this fact is common knowledge, then speakers will have to play the Gricean strategy. We already know that speakers of type 1 must play the Gricean strategy (since they have no choice), so all we need to show is that speakers of types 2 and 3 must play the Gricean strategy.

First I will argue that it is common knowledge that speakers of type 3 will play the Gricean strategy (and hence that they will actually play it). The argument goes as follows: we assume that it is common knowledge that speakers have not agreed ahead of time to play any particular strategy. We will assume, thus, that a speaker of type 3 will prefer a move m to a move m' if m is guaranteed to yield a payoff u given just the structure of the game and the rationality of the players, while m' never yields a payoff higher than u and is not guaranteed to yield u on those same assumptions alone. Now in this case uttering (6) guarantees the top payoff for a speaker of type 3 no matter what the audience believes about the speaker's strategy, but no other utterance guarantees it in this way. Thus we can reason that given the common knowledge that the players have not agreed to play the game in any particular way, it follows that speakers of type 3 will utter (6) (and thus play the Gricean strategy). When something follows from common knowledge, it is itself be part of common knowledge, so this is also common knowledge.

Now, I will argue that if it is common knowledge that speakers of type 3 play the Gricean strategy, the only rational strategy for speakers of type 2 is to utter (5). This actually follows quite quickly. If a speaker of type 2 utters (5) then, given this assumption, he will induce a belief of probability 1 in his being of type 2 in the audience, and thus get the highest payoff. He cannot do so if he utters (4), so he must utter (5).

We can see, then, that common knowledge of lack of coordination makes one strategy compelling. The speaking situation we are considering is clearly one in which there is such lack of coordination (after all, the semantic conventions are meant to exhaust the pre-arranged aspects of the game), so the Gricean strategy really is rationally compelling in these cases. In the next section, I will generalize this style of reasoning to a wider class of games. Readers uninterested in the details should skip this section.[13]

[13] van Rooij (to appear) also considers three-point scales. However, he decides for the Gricean strategy in part by simply stipulating that strategies such as the Partial Pooling one are not allowed.

3.2 A Generalization

In this section, I formalize an idea of how people should play games when it is common knowledge that there is no prior coordination on how to play (in other words, it is commonly known that there are no pre-established conventions of play). I then prove that this notion will show that speaker's must choose Gricean strategies in the types of games we have considered above.

I will not restrict myself to considering the particular scalar-choice signaling games we looked at above but instead consider a wider class of games. The formalization is cumbersome, but I think it respects our intuitions of how people should act in these situations. First, some definitions:

Conventionlessly Dominates A move m conventionlessly dominates a move m' if for some n, m is conventionlessly(n) preferred to m'. A move m is conventionlessly dominant if it is not conventionlessly dominated by any other move. A strategy s conventionlessly dominates s' if s contains a move m and s' contains a move m' s.t. m conventionlessly dominates m' and s' doesn't have a move that conventionlessly dominates a move in s.

Conventionlessly(0) Dominates A move m *conventionlessly(0) dominates* a move m' iff there exists a payoff u that satisfies these conditions.

1. It follows from the common knowledge of the structure of the game and common knowledge that both players are rational and that m yields a payoff of at least u (for each player).
2. It does not follow from the common knowledge of the structure of the game and common knowledge that both players are rational and that m' will give u.

3. m' never yields a payoff higher than u

Conventionlessly(n) Dominates For $n > 0$, a move m is *conventionlessly(n) dominates* a move m' if 1 or 2 obtains:

1. m conventionlessly($n-1$) dominates m'.
2. There exists a payoff u that satisfies these three conditions:
   (a) It follows from common knowledge of the structure of the game, common knowledge that both players are rational, and common knowledge that players will choose conventionlessly($n-1$) dominant moves that m yields a payoff of at least u (for each player)
   (b) It does not follow from common knowledge of the structure of the game, common knowledge that both players are rational, and common knowledge that players will choose conventionlessly($n-1$) dominant moves that m' will yield u.
   (c) m' never gives a payoff higher than u.

I think that in the case of common knowledge of lack of coordination on any particular strategy, players should not play conventionlessly-dominated strategies. This is because players can each assume that each knows that no one will play a conventionlessly(0) dominated strategy since, if there is lack of coordination, such a strategy is less "safe" than the strategy that dominates it. Moreover each player knows they each knows this, and thus can reason similarly that no one will play a conventionlessly(1) strategy, and so on.... So this is not just an arbitrary equilibrium concept, but one that should manifestly govern the actions of people in this sort of situation.

Now let us consider certain two-player signaling games with n-types of speaker that have the feature that for a speaker of a given type, the highest payoff for both players can be achieved if the audience has completely accurate beliefs about the speaker's type after the speaker sends his message (this includes not only my kind of game, but also many more standard cooperative signaling games). We need to make two further assumptions about the structure of the game: For each type of speaker, i, there is a unique message k available to him such that no speaker of type $< i$ can also use k. And one piece of terminology: call the *Gricean strategy* in this game the strategy on which each a speaker of type k sends the unique message l s.t. no player of type $i < k$ can also send l. Call games of this type *well-structured communication games.* (It should be clear that the two-point and three-point scale games, considered above, are examples of well-structured communication games, in addition to certain other games with non-linear structures.)

I will now give a proof by induction that the Gricean strategy conventionlessly dominates all others in well-structured communication .

Fact 1. *In a well-structured communication game with n players, the Gricean strategy is conventionlessly dominant.*

Proof. First the base case:

Fact 2. *For speakers of type n the conventionlessly(0) dominant move is the Gricean one.*

Proof. The Gricean move is uniquely available to speakers of type n, so audiences who receive this message will have to assume that the speaker is of type n with probability 1. For any other move, the audience could rationally assign a probability of less than 1 to the proposition that the speaker is of a different type. So the Gricean strategy is conventionlessly(0) dominant for players of type n. □

Now the induction step:

Fact 3. *If for all players of type $> n - i$ the Gricean strategy is conventionlessly($i - 1$) dominant over all others, then for players of type $n - i$ the Gricean strategy will be conventionlessly(i) dominant over all others.*

Proof. We just need to prove that given common knowledge that all players will play conventionlessly($i - 1$) dominated strategy, then it follows that the Gricean strategy is the only strategy that yields the correct belief in the audience given just this common knowledge assumption along with the common knowledge of the structure of the game and the player's rationality. Given the common knowledge assumption the only sender who can play the Gricean move for type $n - i$ is a player of type $n - i$, so the audience will have to believe with credence 1 that the player is of type $n - i$ if the speaker makes the Gricean move for his type. For all other available moves it is clear that this does not follow, since it is consistent with the assumptions we have made that other senders of lower types may also use those messages. $\square$

4 A Different Kind of Case

Our previous cases have all been very simple. It was a bit surprising how hard it is to show that Gricean speaker strategies are rationally compelling in such very simple cases. However, what is perhaps even more surprising, is that Gricean speaker strategies are probably *not* rationally compelling in slightly more complex cases.

The case I will consider is one much discussed in the recent literature on implicatures, though in a slightly different context. This is the case where there are *two* scalar terms in one sentence. The recent literature on these sorts of cases has raised the question of whether the Gricean programme can handle such cases at all.[14]

[14] Chierchia (2004) is largely responsible for the recent interest in this topic. He argues that scalar

Let me consider one simple case and show that the sort of Gricean reasoning I outlined at the beginning of this paper can plausibly be extended to cover this particular case. Consider an utterance of (7):

(7) Some of the students answered some of the questions.

Let us assume that there was a salient choice between "some" and "all' for both instances of "some" in this sentence. Let us also assume that these choices were made simultaneously. In this case the speaker was choosing between four sentences. The one he uttered, (7), and these three others:

(8) All of the students answered some of questions.

(9) Some of the students answered all of the questions.

(10) All of the students answered all of the questions.

Now we assume that a speaker utters the most informative relevant sentence that he can (the maxim of quantity). Thus (assuming the relevance of all these utterances) if the speaker uttered (7) it is because he did not believe (8)–(10). It's debatable if this result matches what we find empirically (which depends so much on focal intonation), but for these purposes let's assume it does.

Can we derive these implicatures on grounds independent of the maxim of quantity? It might seem, from previous discussion that would should be able to, but a little thought can show us that a rather thorny issues arises.

First we need to describe the situation more precisely. Let's assume again that the speaker is trying to maximize the information he conveys *by whatever means* to the

implicatures need to be embedded compositionally into the semantics. Gricean responses to Chierchia's problems are given in Sauerland (2004), Schulz and van Rooij (2006), and Spector (2006).

audience. Let's also assume that the only means to convey information (besides direct assertion) is by some implicature about the extent of the speaker's beliefs based on what he uttered. Note first that there are actually *five* relevant states of belief[15] for the speaker with respect to these *four* alternatives sentences (I'll label them with letters so it's easier to keep track of):

Type *ss* Believes only (7)

Type *as* Believes only (7) and (8)

Type *sa* Believes only (7) and (9)

Type *as/sa* Believes only (7), (8) and (9)

Type *aa* Believes (7), (8), (9), and (10)

Here is a version of the intuitive problem: Suppose we *assume* that a speaker of type *as* says (8), a speaker of type *sa* says (9), and a speaker of type *aa* says (10) . These assumptions are, after all, things we can derive from the maxim of quantity. If we assume these facts, we *cannot* derive as a further piece of Gricean behavior, that a speaker of type *as/sa* should utter either (8) or (9). The reason we cannot derive this is that all we are assuming is that a speaker will want to convey as much as possible by his utterance. However, given our assumptions, an utterance of (7) will convey that the speaker is not of type *sa*, *as*, or *aa* whereas an utterance of type (8) will convey that the speaker is not of type *ss*, *sa* or *aa* and an utterance of (9) will convey that the speaker is not of type *ss*, *as* or *aa*. None of these utterances is strictly more informative than the any other, so, given our assumptions, there is no way to derive the result that a speaker of type *as/sa* should say (8) or (9) rather than (7). The problem is this: once we assume

[15] Again I'm excluding negative beliefs, i.e. beliefs in the falsity of these sentences, since I assume that inferences about these come indirectly as consequences about inferences about positive belief states (or their lack) and speaker expertise.

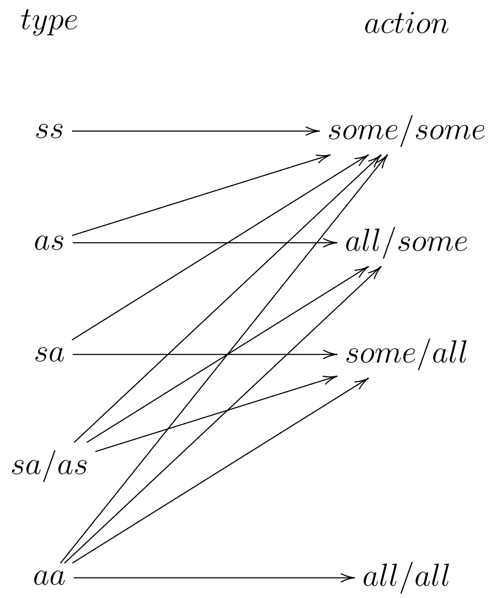

Figure 5: Speaker options in the simultaneous-choice game

the speaker follows the maxim of quantity in certain cases, new implicatures arise which prevent us from showing that he needs to say the strongest statement he believes in other cases.

This is just to show that the whole Gricean strategy is not even compelling when we assume that speakers are following almost all of it. Whether it is compelling without any assumptions is harder to see. At the least, this indicates that we need a more formal approach to assessing what strategy the rational speaker (and hence the audience) should choose.

So let us again look at the signaling game associated with this utterance choice, which is in Figure 5. We will assume in this game, again, that the payoffs are proportional to the credence the audience puts in the speaker being of the type that he actually is.

Here is a version of the Gricean strategy for this game (see Figure 6):

Gricean Type *aa* utters (7), type *as* utters (8), type *sa* utters (9), type *as*/*sa* utters (8) or (9) (say with 50/50 probability), and type *aa* utters (10).

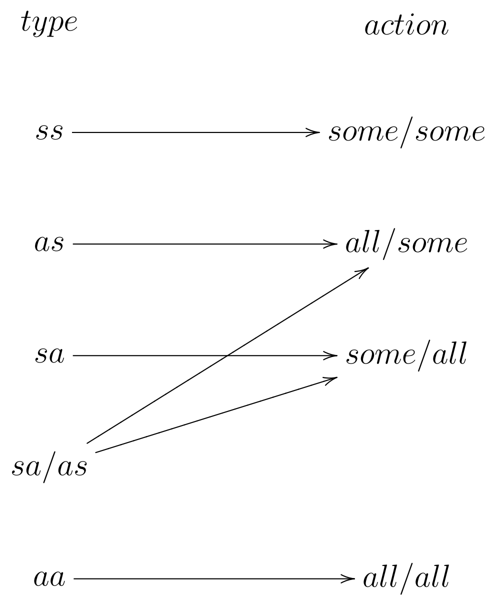

Figure 6: Gricean strategy in simultaneous-choice game

This strategy is obviously a PBE. As with the three-scale game it is provable that any strategy on which players of type *aa* do not utter (10) will be untenable if it is believed by the speaker that the audience is not completely certain as to whether the players are playing a non-Gricean strategy. However, there are, in this case, non-Gricean PBEs on which it is common knowledge that speakers of type *aa* utter (10). Here is one such strategy, which is in some ways more natural than any Gricean strategy since it is pure as well asymmetric[16] (see Figure 7):

Pure Type *aa* utters (7), type *as* utters (8), type *sa* utters (9), type *as*/*sa* utters (7), and type *aa* utters (10).

These two strategies are both, I think, *entirely* reasonable. Neither is Pareto-dominant over the other. Nor does either equilibrium have the feature that we saw in the three-stage game where one has parts which are rationally compelled in certain cases of uncer-

[16] By symmetric, I mean that the symmetries in the structure of the game are mirrored in the strategies. The Gricean strategy is symmetric, but mixed.

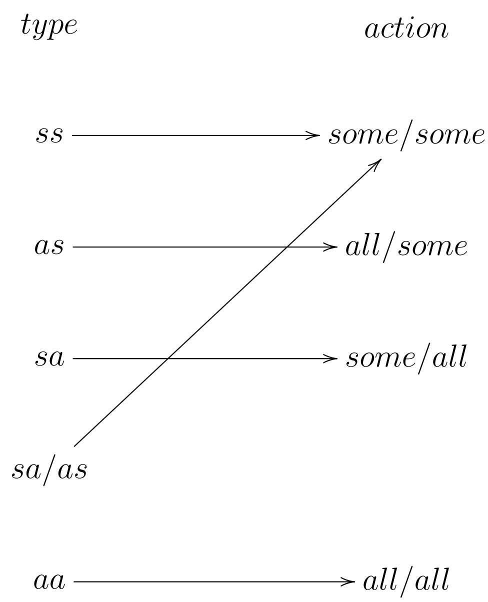

Figure 7: Pure strategy in simultaneous-choice game

tainty (the game is obviously not a well-structured communication game in the sense I defined in Section 3.2). Obviously, the Gricean strategy has certain properties and (perhaps) advantages that the non-Gricean strategy does not have. Nonetheless, I think it would be a hopeless task to argue that one of the two strategies is *rationally compelling* for speakers.

As the informal reasoning I gave above suggests, the case of choosing two scalar terms simultaneously is very complex. The complications presumably arises from the fact that there are more relevant speaker information states than there are utterances, thus there is more than one possible way to use one utterance to cover different information types while still using all the utterances.[17] My conclusion is that, in these cases, Gricean behavior according to the maxim of quantity cannot be independently grounded in a desire to be as informative as possible. Making the strongest possible statement is only one of many strategies for being maximally informative. Unfortunately, without

[17]There is a brief discussion in van Rooij (to appear) of cases where there are more speaker types than utterances.

coordination, none of these strategies is guaranteed to work. If the Gricean account is right, then, we need to have maxims in place as *conventions* to solve this coordination problem.

Note that if we add more alternatives to our choice situation we will no longer derive these same results. However, given the symmetry problem it is clear that we need to sharply restrict which alternatives are available to a given speaker. My point here is only to show that on the most standard assumptions about what the alternatives are in this case, the Gricean strategy does not follow from more basic assumptions about rationality.

5 Conclusion

This is a very limited study of the degree to which Gricean derivations can be derived without stipulating maxims as primitive conventions. Many simple cases do seem to not require the maxims to be stipulated. In more complex cases we either need to stipulate maxims or make assumptions about the structure of alternatives that may be quite hard to motivate. My aim here is not to advocate a particular way of thinking about implicatures, but rather to show that we need to be wary about taking Gricean explanations as defaults across all cases.

References

Benz, Anton, Jager, Gerhard, and van Rooij, Robert (eds.) (2005). *Game Theory and Pragmatics*. Palgrave MacMillan.

Chierchia, Gennaro (2004). Scalar implicatures, polarity phenomenon, and the syntax/pragmatic interface. In Andrea Belleti (ed.), *Structures and Beyond*. Oxford University Press.

Fox, Danny (2006). Free choice and the theory of scalar implicatures. Unpublished manuscript.

Gibbons, Robert (1992). *Game Theory for Applied Economists*. Princeton University Press.

Horn, Larry (1972). *The Semantics of the Logical Operators in English*. Ph.D. thesis, UCLA.

de Jager, Tikitu and van Rooij, Robert (2007). Explaining quantity implicatures. In *Proceedings of the 11th conference on theoretical aspects of rationality and knowledge*.

Kroch, Anthony (1972). Lexical and inferred meanings from some time adverbs. *Quarterly Progress Report of the Research Laboratory of Electronics*, 104.

Lewis, David (1969). *Convention*. Harvard University Press.

van Rooij, Robert (to appear). Games and quantity implicatures. *Journal of Economic Methodology*.

Sauerland, Uli (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27:367–391.

Schulz, Katrin and van Rooij, Robert (2006). Pragmatic meaning and non-monotonic reasoning. *Linguistics and Philosophy*, 29(2):205–250.

Spector, Benjamin (2006). *Aspects de la pragmatique des opérateurs logiques*. Ph.D. thesis, Universite Paris 7, Denis Diderot.