Cooperative versus argumentative communication

Robert van Rooy*
Institute for Logic, Language and Computation
University of Amsterdam
R.A.M.vanRooij@uva.nl

Abstract

Game theoretical analyses of communication (e.g. Lewis, Crawford & Sobel) demand cooperation between conversational partners for reliable information exchange to take place. Similarly, in pragmatics, the theory of language use, it is standard to assume that communication is a cooperative affair. Recently, this standard view has come under attack by Durcot and Merin, and it has been proposed that an argumentative view on natural language use is more appropriate. In this paper I discuss to what extent this attack is justified and whether the alternative view can provide a more adequate analysis of 'pragmatic meaning', i.e., implicatures. I will investigate the game-theoretical underpinning of the argumentative view, and contrast Merin's analysis of scalar implicatures with one using the principle of exhaustive interpretation.

1 Introduction

Language is one of the most precious gifts history has bestowed upon us. It allows us to cooperate better with other individuals by enabling us (among others) to communicate useful and relevant information. Indeed, many (e.g. Maynard Smith & Szathmary, 1999) see the introduction of language as one of the major transitions in the evolution of life exactly for this reason. But not only can the use of language enhance cooperation, it is widely assumed in linguistics (e.g. Grice 1967, Lewis 1969) that cooperation is already required for communication to take place in the first place. Recently, however, this standard view on communication has come under attack. In this paper I discuss to what extent this attack is justified.

After introducing communication games as discussed in Lewis' (1969) classical work on conventions, I give a short overview of some recent work on communication by economists. I discuss in particular some work that shows in which circumstances how much communication is possible. In section 3, I discuss the arguments of Oswald Ducrot and Arthur Merin that language use is an argumentative, instead of a cooperative affair. I will focus on their analysis of adversary connectives and of scalar implicatures. Merin's view on communication as a bargaining game plays a crucial role here. In part 4, I discuss whether the scalar implicatures discussed by Ducrot and Merin are really problematic for a Gricean view on communication. Section 5 concludes the paper.

*The research of this work is supported by a fellowship from the Royal Netherlands Academy of Arts and Sciences (KNAW), which is gratefully acknowledged. For writing this paper I profited a lot from discussions with Katrin Schulz on exhaustive interpretation and Merin's analysis of implicatures. Thanks also to Martin Stokhof for discussion.

2 Communication in games

2.1 Cheap talk games

In his classic work on conventions, Lewis (1969) proposed to study communication by means of so-called signaling games. In this section we will only consider cheap talk games: games where the messages are not directly payoff relevant. A signaling game with payoff irrelevant messages is a sequential game of incomplete information with two players involved, player 1, the sender, and player 2, the receiver. Both players are in a particular state, an element of some set T. Player 1 can observe the true state, but player 2 can not. The latter has, however, beliefs about what the true state is, and it is common knowledge between the players that this belief is represented by probability function P over T. Then, player 1 observes the true state t and chooses a message m from some set M. After player 2 observes m (but not t), he chooses some action a from a set A, which ends the game. The utilities of both players are given by $U_1(t, a)$ and $U_2(t, k)$. The (pure) strategies of the player 1 and player 2 are elements of $[T \to M]$ and $[M \to A]$, respectively. In simple communication games, we call these functions sending and receiving strategies, i.e., S and R.

What strategy combinations are *equilibria* of the game depends on the probability distribution. With distribution P, the strategy pair $\langle S, R\rangle$ is an equilibrium if, as usual, neither player can do any better by unilateral deviation. First, for each t in T, the player 1's message $S(t)$ must maximize her utility, given that player 2 is using strategy R. Thus, for each $t \in T$, $S(t)$ must solve $max_{m\in M} U_1(t, R(m))$. Also player 2 must maximize his utility, but to determine this we have to take into account what he believes after he received a message. It is assumed that after observing any message m from M, the probability distribution representing player 2's belief about which situation player 1 could have been in when sending m is denoted by $\mu(t/m)$, and defined in terms of P as follows: $\mu(t/m) = \frac{P(t)}{\sum_{t\in S^{-1}(m)} P(t)}$. In this definition $S^{-1}(m)$ denotes the set of situations in which player 1, by using strategy S, sends m. Now we can determine what player 2 has to do to maximize his utility given that player 1 uses strategy S: for each message m, he has to solve $max_{a\in A}\mu(t/m) \times U_2(t, a)$.

In cheap talk games, the messages are not directly payoff relevant: the utility functions do not mention the messages being used. Thus, the only effect that a message can have in these games is through its information content: by changing the receiver's belief about the situation the sender (and receiver) is in. If a message can change the receiver's beliefs about the actual situation, it might also change the receiver's optimal action, and thus indirectly affect both players' payoffs.

In different equilibria of a cheap talk communication game, different amounts of information can be transmitted. But for cheap talk to allow for informative communication at all, a sender must have different preferences over the receiver's actions when he is in different states. Likewise, the receiver must prefer different actions depending on the what the actual situation is (talk is useless if the receiver's preferences over actions are independent of what the actual situation is.) Finally, the receiver's preferences over actions must not be completely opposed to that of the sender's. But how informative can cheap talk be? That is, how fine-grained can and will the sender reveal the true situation if talk is cheap?

In an important article, Crawford & Sobel (1982) show that the amount of credible communication in these games depends on how far the preferences of the participants are aligned. To illustrate, assume that the state, message and action spaces are continuous and between the interval of zero and one. Thus, $T = [0, 1]$; the message space is the type space ($M = T$) and also the action space is in the interval $[0, 1]$. Now, following Gibson (1992), we can construct as a special case of their model the following quadratic utility functions for sender and receiver such that there is a single parameter, $b > 0$, that measures how closely the preferences of the two players are aligned:

$$\begin{aligned} U_2(t,a) &= -(a-t)^2 \\ U_1(t,a) &= -[a-(t+b)]^2 \end{aligned}$$

Now, when the actual situation is t, the receiver's optimal action is $a = t$, but the sender's optimal action is $a = t + b$. Thus, in different situations the sender has different preferences over the receiver's actions (in 'higher' situations senders prefer higher actions), and the interests of the players are more aligned in case b comes closer to 0. Crawford & Sobel (1982) show that in such games all equilibria are *partition equilibria*; i.e., the set of situations T can be partitioned into a finite number of intervals such that senders in a state belonging to the same interval send a common message and receive the same action. Moreover, they show that the amount of information revealed in equilibrium increases as the preferences of the sender and the receiver are more aligned. That is, the closer parameter b approaches 0, there exists an equilibrium where the sender will tell more precisely in which situation she is in, and thus more communication is possible. However, when parameter b has the value 1, it represents the fact that the preferences of sender and receiver are opposed. A sender in situation $t = 1$, for instance prefers most action $a = 1$, and mostly disprefers action $a = 0$. If $b = 1$, however, a receiver will prefer most action $a = 0$ and most dislikes action $a = 1$. As a result, no true information exchange will take place if $b = 1$, i.e., if the preferences are completely opposed.

To establish the fact proved by Crawford & Sobel, no mention was made of any externally given meaning associated with the messages. But, in equilibrium the players do associate a particular meaning with the messages: If $\langle S, R\rangle$ is commonly known to be the equilibrium used by the two players, they commonly know that the meaning of any message m is $S^{-1}(m)$ (if the message was used indicatively) or $R(m)$ (if the message was used imperatively). Now assume, in addition, that the messages *do* have an externally given meaning, and take this meaning to be a set of situations. Thus, let us assume an externally given interpretation function $[\cdot]$ that assigns to every $m \in M$ a subset of T. As it turns out, this doesn't have any effect on the equilibria of the game, and on Crawford & Sobel's (1982) result on the amount of possible communication in cheap talk games. At least, if the games are still taken to be ones of cheap talk: if no constraints are imposed on what kinds of meanings can be expressed in particular situations, for instance if no requirements like $\forall t \in T : t \in [S(t)]$, saying that speakers have to tell the truth, are put upon speakers' strategies S.

2.2 Verifiable communication with sceptic audience

What happens if we *do* put extra constraints upon what can and what cannot be said? As it turns out, this opens up many new possibilities of credible communication. In fact, Lipmann & Seppi (1995) (summarized in Lipman, 2002) have shown that with such extra constraints, interesting forms of reliable information transmission can be predicted in games where you expect it the least: in debates between agents with opposing preferences.

Before we will look at debates, however, let us first consider cheap talk games when we assume that the signals used come with a pre-existing meaning and, moreover, that speakers always tell the truth. This still doesn't guarantee that language cannot be used to mislead one's audience. Consider a two-situation two-action game with the following utility table.

two-situation, two-action:

	a_H	a_L
t_H	$1,1$	$0,0$
t_L	$1,0$	$0,1$

In this game, the informed sender prefers, irrespective of the situation she is in, column player to choose a_H, while column player wants to play a_H if and only if the sender is in situation t_H. Now assume that the expected utility for player 2 to perform a_H is higher than that of a_L (because $P(t_H) > P(t_L)$), and suppose that we demand truth: $t_i \in [[S(t_i)]]$. In that case, the rational message for an individual in the 'high' situation to send is one that conventionally expresses $\{t_H\}$, while an individual in the 'low' situation has an incentive to send a message with meaning $\{t_H, t_L\}$. If the receiver is naive he will choose a_H after hearing the signal that expresses $\{t_H, t_L\}$, because a_H has the highest expected utility. A more sceptical receiver, however, will argue that a speaker that sends a message with meaning $\{t_H, t_L\}$ must be one that is in a 'low' situation, because otherwise the speaker could, and thus should (in her own best interest) have sent a message with meaning $\{t_H\}$. Thus, this sceptical receiver will reason that the sender was in fact in a low-type situation and interprets the message as $\{t_L\}$. In general, suppose that the sender has the following preference relation over a set of 10 situations: $t_1 < t_2 < ... < t_{10}$ (meaning that t_1 is the worst situation) and sends a message m with pre-existing meaning $[m]$. A sceptical receiever would then assign to m the following pragmatic interpretation $S(m)$ based on the speaker's preference relation '$<$', on the assumption that the speaker knows which situation she is in:

$$S(m) \quad = \quad \{t \in [m] | \neg\exists t' \in [m] : t' < t\}$$

This pragmatic interpretation rule is based on the assumption that the speaker gives as much information as she can that is useful to her, and that the hearer anticipates this speaker's maxim (to be only unspecific with respect to more desirable states) by being sceptical when the speaker gives a message with a relatively uninformative meaning.[1]

Now consider debates in which the preferences of the participants are mutually opposed. Suppose that players 1 and 2 are two such debaters who *both* know the true state. Now, however, there is also a third person, the *observer*, who doesn't. Both debaters present evidence to the observer, who then chooses an action $a \in A$ which affects the payoffs of all of them. We assume that the observer's optimal action depends on the state, but that the preferences of the debaters do not. In fact, we assume that the preferences of players 1 and 2 are strictly opposed: in particular, if player 1 prefers state t_i above all others, t_i is also the state that player 2 dislikes most. By assuming that the utility functions of all three participants are of type $U_j(t, a)$, we again assume that the message being used is not directly payoff relevant, just as in cheap talk games.

We assume that each debater can send a message. Just as before, let us assume that S denotes the strategy of player 1, and R the strategy of player 2. In contrast to the cheap talk games discussed above, we now crucially assume that the messages have an externally given meaning given by interpretation function $[\cdot]$. Let us first assume that while debater 1 can make very precise statements, i.e., that a particular state t holds, debater 2 can only make very uninformative statements saying that a particular state is *not* the case. Let us assume for concreteness that $T = \{t_1, ..., t_{10}\}$. Then the 'meaning' of $S(t_i)$, $[S(t_i)]$ can consist of one state, $\{t_j\}$, while the meaning of $R(t_i)$, $[R(t_i)]$, always consists of 9 states. Thus, debater 1 can be much more informative about the true state, and is thus in the advantage. But player 2 has an advantage over player 1 as well: in contrast to for debater 1, it is commonly known of debater 2 that he is reliable and will only make *true* statements. Thus, for all $t_i \in T : t_i \in [R(t_i)]$, while it might be that $t_i \notin [S(t_i)]$.

Suppose now that the observer may ask two statements of the players. The question is, how much information can the observer acquire? One is tempted to think

[1] If hearers use such an interpretation rule, speaker have no reason anymore to be vague. But, of course, vagueness can still have positive pay-off when one's audience is unsure about your preferences.

that the messages cannot really give a lot of information: Debater 1 has no incentive to tell the truth, so acquiring two messages from him is completely uninformative. Debater 2 will provide true information, but the informative value of her messages is very low: after two messages from him the observer still doesn't know which of the remaining 8 states is the true one. Surprisingly enough, however, the observer can organize the debate such that after two rounds of communication, he knows for certain which state actually obtains.

The trick is the following: the observer first promises, or warns, debater 1 that in case he finds out that the latter will not give a truthful message, he will punish debater 1 by choosing the action that is worst for her. This is possible because it is common knowledge what the agents prefer. For concreteness, assume that debater 1 has the following preferences $t_{10} > t_9 > ... > t_1$. Afterwards, the observer first asks debater 1 which state holds, and then asks debater 2 to make a statement. Suppose that the first debater makes a very informative statement of the form 'State t_i is the true state'. Obviously, debater 2 will refute this claim, if it is false. For in that case the observer will as a result choose the state most unfavorable to debater 1, and thus most favorable to debater 2, i.e. t_1. Thus, if she is precise, debater 1 has an incentive to tell the truth state, and the observer will thus learn exactly which state is the true one. Suppose that the true state is the one most undesirable for player 1, t_1. So, or so it seems, she has every reason to be vague. Assume that debater 1 makes a vague statement with meaning $\{t_i, ..., t_n\}$. But being vague now doesn't help: if the true state is ruled out by this vague meaning, debater 2 will claim that (even) the least preferred state in it is not true, and if debater 2 doesn't refute debater's 1 claim in this way the observer will choose the most unfavorable state for debater 1 compatible with the true message with meaning $\{t_i, ..., t_n\}$. In general, if debater's 1 message m has meaning $[m]$, and if m is not refuted, then the observer will 'pragmatically' interpret m as follows: $\{t \in [m] | \neg\exists t' \in [m] : t' < t\}$, where '$t' < t$' means that debater 1 (strictly) prefers t to t'. Notice that this is exactly the pragmatic interpretation rule $S(m)$ described above. From a signaling game perspective, this just means that the game has a completely separating equilibrium: whatever the true state is, it is never in the interest of debater 1 not to say that this is indeed the true state.

The example discussed above is but a simple special case of circumstances characterized by Lipman & Seppi (1995) in which observers can 'force' debaters to provide precise and adequate information, even though they have mutually conflicting preferences. Glazier & Rubinstein (2001) have generalized this by studying optimal debate procedures in circumstances in which full information revelation cannot be forced upon speakers.[2]

The discussion in this section shows that truthful information transmission is possible in situations in which the preferences of the conversational participants are mutually opposed. This seems to be in direct conflict with the conclusion reached in section 2.1, where it was stated that credible information transmission is impossible in such circumstances. However, this conflict is not real: On the one hand, a central assumption in cheap talk games is that talk is really cheap: one can say what one wants because the messages are *not verifiable.* The possibility of credible information transmission in debates, on the other hand, crucially depends on the assumption that claims of speakers are verifiable to at least some extent, i.e., they are falsifiable (by the second debater), and that outside observers can punish the making of misleading statements. In fact, by taking the possibility of falsification and punishment into account as well, we predict truthful communication also in debates, because the preferences of the agents which seemed to be opposing are still very much aligned at a 'deeper' level.[3] This section also shows that if a hearer knows the preferences of the

[2]This work is summarized in chapter 3 of Rubinstein (2000).

[3]In van Rooy (2003, 2004) a similar interpretation is given to the so-called 'costly signaling games' used frequently in economics and biology.

speaker and takes him to be well-informed, there exists a natural 'pragmatic' way to interpret the speaker's message which has already a pre-existing 'semantic' meaning, based on the assumption that speakers are rational and only unspecific, or vague, with respect to situations that are more desirable for them.

3 Communication as a debate

It is well established that a speaker in a typical conversational situation communicates more by the use of a sentence than just its conventional truth conditional meaning. Truth conditional meaning is enriched with what is *conversationally implicated* by the use of a sentence. In pragmatics – the study of language use – it is standard to assume that this way of enriching conventional meaning is possible because we assume speakers to conform to Grice's (1967) *cooperative principle*, the principle that assumes speakers to be rational cooperative language users. This view on language use suggests that the paradigmatic discourse situation is one of cooperative information exchange.

Merin (1999b) has recently argued that this view is false, and hypothesized that discourse situations are paradigmatically ones of explicit or tacit debate. He bases this hypothesis on the work of Ducrot (1973) and Anscombre & Ducrot (1983) where it is strongly suggested that some phenomena troublesome for Gricean pragmatics can be analyzed more successfully when we assume language users to have an *argumentative* orientation. In section 3.1 I will sketch some of Ducrot's arguments for such an alternative view on communication, and section 3.2 will be used to describe Merin's analysis of some implicatures which are taken to be troublesome for a cooperative view on language use.

3.1 Ducrot on rhetorical contrast and scalar reasoning

Adversary connectives The connective *but* is standardly assumed to have the same truth-conditional meaning as *and*. Obviously, they are used differently. This difference is accounted for within pragmatics. It is normally claimed that 'a and b' and 'a but b' give rise to different *conventional implicatures*, or appropriateness conditions. On the basis of sentences like (1) it is normally (e.g. Frege, 1918) assumed that sentences of the form 'a but b' are appropriate, if b is unexpected given a.

(1) John is tall but no good at baseball.

This, however, cannot be enough: it cannot explain why the following sentence is inappropriate:

(2) John walks but today I won the jackpot

Neither can it explain why the following sentence is okay, because expensive restaurants are normally good.

(3) This restaurant is expensive, but good.

Ducrot (1973) and Anscombre & Ducrot (1983) argue that sentences of the form 'a but b' are always used argumentatively, where a and b are arguments for complementary conclusions: they are contrastive in a rethorical sense. For instance, the first and second conjunct of (3) argue in favor of not going and going to the restaurant, respectively.

Not only sentences with *but*, but also other constructions can be used to express a relation of rhetorical contrast (cf. Horn, 1991). These include complex sentences with *while*, *even if*, or *may* in the first clause (i.e. concession), and/or *still*, *at least*, or *nonetheless* either in place of or in addition to the *but* of the second clause (i.e. affirmation):

(4) a. While she won by a {small, * large} margin, she did win.

b. Even if I have only three friends, at least I have three.

c. He may be a professor, he is still an idiot.

Rhetorical contrast is not all to the appropriateness of sentences like (3) or (4). It should also be the case that the second conjunct should be an argument in favor of conclusion h that the speaker wants to argue for. And, if possible, it should be a stronger argument for h than the first conjunct is for $\neg h$. In this way it can also be explained why (3), for instance, can naturally be followed by $h =$ 'You should go to that restaurant', while this is not a good continuation of

(5) This restaurant is good, but expensive.

which is most naturally followed by $\neg h =$ 'You should not go to that restaurant'.

There is another problem for a standard Gricean approach to connectives like *but* that can be solved by taking an argumentative perspective (cf. Horn, 1991). It seems a natural rule of cooperative conversation not to use a conjunctive sentence where the second conjunct is entailed by the first. Thus, (6) is inappropriate.

(6) *She won by a small margin, and winning she did.

However, even though *but* is standardly assumed to have the same truth-conditional meaning as *and*, if we substitute *and* in (6) by *but*, the sentence becomes perfectly acceptable:

(7) She won by a small margin, but winning she did.

If – as assumed by standard Gricean pragmatics – only truth-conditional meaning is taken as input for pragmatic reasoning, it is not easy to see how this contrast can be accounted for. By adopting Ducrot's hypothesis that in contrast to (6) the conjuncts in (7) have to be rhetorically opposed, the distinction between the two examples can be explained easily: if a speaker is engaged in a debate with somebody who argued that Mrs. X has a relative lack of popular mandate, she can use (7), but not (6).

Thus, it appears that an argumentative view on language use can account for certain linguistic facts for which a non-argumentative view seems problematic. But how then can we explain the fact that in many appropriate uses of 'a but b', the second conjunct is unexpected given the first? As we will see soon, this can naturally be explained as a special case of Durcot & Anscombre's appropriateness condition, given Merin's (1999a) formalization.

Scalar reasoning Anscombre & Ducrot (1983) argue that to account for so-called 'scalar implicatures' an argumentative view is required as well. Scalar implicatures are normally claimed to be based on Grice's maxim of *quantity*: the requirement to say as much as one can (about a topic of conversation). This gives rise to the principle that everything 'higher' on a scale than what is said is false, where the ordering on the scales is defined in terms of informativity. Standardly, scales are taken to be of the form $\langle P(k), ..., P(m)\rangle$, where P is a simple predicate (e.g. *Mary has x children*) and for each $P(i)$ higher on the scale than $P(j)$, the former must be more informative than the latter.[4] From the assertion that $P(j)$ is true we then conclude by scalar implicature that $P(i)$ is false. For instance, if Mary says that she has two children, we (by default) conclude that she doesn't have tree children, because otherwise she could and should have said so (if the number of children is under discussion). Other

[4]See Parikh (2001) for an informal game-theoretical derivation of scalar implicatures.

examples are scales like $\langle a \wedge b, a \vee b\rangle$: from the claim that John *or* Mary will come, we are normally allowed to conclude that they do not come both.

Unfortunately, as observed by Fauconnier (1975), Hirschberg (1985) and others, we see inferences from what is not said to what is false very similar to the ones above, but where what is concluded to be false is *not more informative* then, or does not entail, what is actually said. For instance, if Mary answers at her job-interview the question whether she speaks French by saying that her husband does, we conclude that she doesn't speak French herself, although this is not semantically entailed by Mary's answer. Such scalar inferences are, according to Anscombre & Ducrot, best accounted for in terms of an argumentative view on language: Mary wants to have the job, and for that it would be more useful that she herself speaks French than that her husband does. The ordering between propositions should not be defined in terms of informativity, or entailment, but rather in terms of argumentative force. Thus, from Mary's claim that her husband speaks French we conclude that the proposition which has a higher argumentative value, i.e., that Mary speaks French herself, is false.

But why then are scalar implicatures typically induced by scales in terms of informativity? Ducrot argues that this is the case because more informative propositions typically have higher argumentative values.

3.2 Merin's analysis of argumentative discourse

Ducrot and Anscombre did not formalize their rhetorico-pragmatic theory of argumentation. This is rather unfortunate, because it leaves their proposal rather vague and difficult to test. Happily for us, Merin (1999a,b) set himself to the task of removing this obstacle and proposes a formalization of their intuitions by making use of the theory of games and decisions. In particular, in Merin (1999b) he proposes to model communication as a *bargaining game* between two agents who have dual preferences with respect to a dichotomic epistemic issue: the question whether a particular proposition h is true or false: If one agent prefers h to be true, the other prefers it to be false. These preferences (together with the beliefs) are used to determine a precise notion of *relevance* of new pieces of information, and he uses this notion to characterize the circumstances in which certain expressions can be used appropriately, and to account for scalar implicatures.

Merin proposes to use Good's (1950) notion of *weight of evidence* as his notion of relevance. Good attributes the notion to Turing, though – according to Merin (1999b) – the notion was already used by Peirce, who called it the *weight of argument*. It is defined with respect to a context represented by a probability function P and a goal proposition h as follows (where $P(a/h)$ denotes the conditional probability of a given h, and is defined as $\frac{P(a \wedge h)}{P(h)}$.):[5]

Definition 1 *(Weight of Argument):* $WA(h, a) = log\frac{P(a/h)}{P(a/\neg h)}$.

Naturally, Merin calls proposition a *positively relevant* to h iff $WA(h, a) > 0$. Similarly, a is called *negatively relevant* and *irrelevant* iff $WA(h, a) < 0$ and $WA(h, a) = 0$, respectively. The fact that informative propositions can be negatively relevant will be important for Merin's analysis of linguistic data. Notice that $W(h, a)$ can equivalently be described as $logP(a/h) - logP(a/\neg h)$, from which we derive immediately the following fact:

Fact 1 $WA(\neg h, a) = -WA(h, a)$.

This fact captures the intuition that if a is a good argument for, or positively relevant to $\neg h$, it is a bad argument for, or negatively relevant to h. Merin (1999a)

[5] With Merin (1999b) we assume that $WA(h, a) = 0$, if $P(h) = 0$ or if $P(h) = 1$, although in these circumstances $P(a/h)$ and $P(a/\neg h)$ are not really defined.

shows that this fact can be used to explain some intuitions Anscombre and Ducrot have about *but*.[6]

Adversary connectives For instance, in analogy with Anscombre & Ducrot, Merin (1999a,b) states that a conjunctive statement of the form 'a but b' is appropriate only if $WA(h,a) < 0$ and $WA(h,b) > 0$ with respect to a contextually given (or derived) goal proposition h. This obviously accounts for (3) 'This restaurant is expensive, but good.'

What makes Merin's (1999a) formalization interesting is the fact that it allows him to explain in a formal rigorous manner why (7) 'She won by a small margin, but winning she did' can be appropriate, although the second conjunct is entailed by the first, and he can derive the appropriateness condition of (1) 'John is tall, but no good at baseball' as a special case: that the second conjunct is unexpected given the first. The possibility of explaining sentences like (7) depends on the fact that even if b is entailed by a, $a \models b$, it is still very well possible that there are h and $WA(\cdot,\cdot)$ such that $WA(h,a) < WA(h,b)$. Thus, the notion of relevance used by Merin does not increase with respect to the entailment relation. In fact, however, Merin (1999a) predicts (7) to be bad, because he also demands that for 'a but b' to be appropriate, it has to be the case that $WA(h, a \wedge b) > 0$. Of course, if $a \models b$, then $WA(h, a \wedge b) = WA(h,a)$, and thus this last condition can never be satisfied for (7) together with $WA(h,a) < 0$. I conclude that Merin (1999a) should have given up this extra requirement as a hard constraint. Thus, perhaps rhetorical contrast *plus* the convention that the second conjunct is a positive argument as far as the speaker is concerned is all there is to adversary connectives.[7] In fact, I don't really believe we need to assume more to account for the natural continuations of 'but'-sentences as discussed in the previous subsection.

To explain the fact that normally the second conjunct of a sentence like 'a but b' is unexpected given the first conjunct, Merin (1999a) makes use of a notion of independence:

Definition 2 *(Independence) a and b are independent conditionally on h and ¬h with respect to P,* $[(a \perp b | \pm h)]_P$ *iff* $P(a \wedge b/h) = P(a/h) \times P(b/h)$ *and* $P(a \wedge b/\neg h) = P(a/\neg h) \times P(b/\neg h)$.

Now, using a result of Reichenbach (1956), Merin (1999a) proves (page 115-16) the following fact (where $sgn(AW(h,a)) = +/-/0$ iff $WA(h,a) > / < / = 0$):

Fact 2 *For any* a, b *and* h*: if* $[(a \perp b | \pm h)]$,
then $0 \neq sgn(WA(h,a)) = -sgn(WA(h,b))$ *iff* $P(b/a) < P(a)$.

Thus, given the independency condition, the unexpectedness condition of b given a follows from Merin's appropriateness condition for sentences of the form 'a but b'. I think this is a very pleasing result. Moreover, Merin proves that unexpectedness of b follows in special circumstances also without making this condition. Take h to be b. Then, obviously, both conditions $WA(h,a) < 0$ and $WA(h,b) > 0$ are satisfied if for example (1) men are expected to be good at baseball, if they are tall. Merin (1999a) shows that if $h = b$, the unexpectedness of b given a, $P(b/a) < P(b)$, follows from his appropriateness condition of 'a but b' stated in terms of the relevance function $WA(\cdot,\cdot)$.

[6]I will limit myself in this paper to (what I take to be) the most essential ingrediants of Merin's (1999a) analysis of *but*.

[7]Though perhaps the requirement of rhetorical contrast is already too strong. I don't know how Merin would account for *Mary came, but I wouldn't have thought so*, an example I owe to Martin Stokhof. Horn (1991) argues that in some cases involving adversary connectives, what is at issue is an opposition between negative and positive *face* of the conjuncts. I agree, and I am not sure to what extent Merin's analysis could account for this.

Scalar reasoning Another interesting feature of adopting function $WA(\cdot, \cdot)$ is that in terms of it we can explain scalar reasoning that cannot be accounted for in terms of the standard assumption that scales have to be ordered in terms of informativity. Consider the example discussed above again, where Mary answers at her job-interview the question whether she speaks French by saying that her husband does. If h is the proposition 'Mary gets the job' – which seems only natural –, her actual answer will have a lower '$WA(h, \cdot)$'-value than the claim that she speaks French herself. Thus, her actual claim, and the claim that she did not make can be ordered in a scale defined by the $WA(h, \cdot)$ values, from which we can derive what is not the case by reasoning with the standard principle that speakers should make their strongest claim possible.

Perhaps somewhat surprisingly, this natural reasoning schema is *not* adopted in Merin (1999b, 2003). In fact, he doesn't want to account for conversational implicatures in terms of the standard principle that everything is false that the speaker didn't say, but could have said (basically, the principle of exhaustive interpretation). Instead, he proposes to *derive* scalar implicatures from the assumption that *all* conversation is a *bargaining game* in which the preferences of the agents are diametrically opposed. From this view on communication, it follows that assertions and concessions have an 'at least' and 'at most' interpretation, respectively:

> if a proponent, Pro, makes a claim, Pro won't object to the respondent, Con, conceding more, i.e. a windfall to Pro, but will mind getting less. Con, in turn, won't mind giving away less than conceded, but will mind giving away more. Put simply: claims are such as to engender intuitions glossable 'at least'; concessions, dually, 'at most'. (Merin, 1999b, p. 191).

This intuition is formalized in terms of Merin's definition of *relevance cones* defined with respect to contexts represented as $\langle P, h\rangle$ (I minimally changed Merin's (1999b) actual definition 8 on page 197.)

Definition 3 The *upward (relevance) cone* ${}^{\geq S}\phi$ of an element ϕ of a subset $S \subseteq F$ of propositions in context $\langle P, h\rangle$ is the union of propositions in S that are at least as relevant to h with respect to P as ϕ is. The *downward (relevance) cone* ${}^{\leq S}\phi$ of ϕ in context $\langle P, h\rangle$ is, dually, the union of S-propositions at most as relevant to h with respect to P as ϕ is.

On the basis of his view of communication as a bargaining game, Merin hypothesizes that while the upward cone of a proposition represents Pro's claim, the downward cone represents Con's default expected compatible counterclaim (i.e., *concession*). Net meaning, then is proposed to be the intersection of Pro's claim and Con's counterclaim: ${}^{\geq S}\phi \cap {}^{\leq S}\phi$, the intersection of what is asserted with what is conversationally implicated.

In the following I will discuss some phenomena concerning scalar reasoning explicitly discussed by Merin (1999b, 2003) where it is claimed that taking an argumentative view on language use has considerable payoff.[8]

Numerals Merin's analysis of conversational implicatures, and the way it differs from the standard Gricean analysis, can perhaps best be illustrated with numerical expressions. In standard Gricean pragmatics it is assumed that numerals have an 'at least'-meaning. The reason why we conclude from the assertion *John has three children* that John has exactly three children is then due to a conversational implicature: the stronger sentence *John has (at least) four children* is entailed to be false. A well-known problem for this analysis is that it also predicts an 'exactly'-reading when it

[8]Other phenomena, like the analysis of free choice inferences, negative polarity items, comparatives, superlatives, and lexical items like *even* and *only* could be discussed as well. However, Merin does not discuss the analysis of these phenomena in enough detail to be summarized confidently.

is explicitly stated that John has *at least* three children, which we don't want. Merin (2003) proposes that numerical expressions have semantically an 'exactly'-meaning, and takes 'at least' to be a modifier. Such an analysis, of course, has to explain the difference in acceptability between the appropriate *John has three children, in fact four* versus the inappropriate *John has four children, in fact three.* Merin proposes to account for this in terms of the notions of upward and downward cones.

Take ϕ to be the proposition that John has three children, and suppose that, for some reason, the speaker wants John to have as many children as possible. Then, the upper cone of ϕ, i.e. ${}^{\geq S}\phi$, should be thought of as the union of propositions of the form *John has n children*, with $n \geq 3$, and thus claims that John has at least 3 children. The downward cone, ${}^{\leq S}\phi$, is now of course the union of propositions of the form *John has n children*, with $n \leq 3$, meaning that John has at most three children. The appropriateness of *John has three children, in fact four* versus the inappropriateness of *John has three children, in fact two* can now be accounted for in terms of whether what is claimed by the first conjunct, ${}^{\geq S}\phi$, is consistent with the meaning of the second conjunct.

The net meaning of a sentence was defined as the intersection of its upward and downward cones: ${}^{\geq S}\phi \cap {}^{\leq S}\phi$. Observe that when ϕ is an element of S, ϕ itself will be a subset of ${}^{\geq S}\phi \cap {}^{\leq S}\phi$ as well. Thus, if $\phi \in S$, the net meaning of ϕ cannot be any stronger than the semantic meaning of ϕ, though it could be weaker![9] Indeed, with respect to normal numerical expressions, Merin's analysis of implicatures is somewhat unexciting. The analysis is still interesting, because for 'at least' sentences it gives better predictions than the standard Gricean analysis does. Notice that although ${}^{\geq S}$[three] $=$ ${}^{\geq S}$[at least three],[10] the downward cones of the two expressions differ: ${}^{\leq S}$[three] $= \{n | n \leq 3\}$, while ${}^{\leq S}$[at least three] $= \{n | \exists m \in$ [at least three] $: n \leq m\}$. In fact, the downward cone of 'At least three men came' is the set of *all* states (of the world). As a consequence, ${}^{\geq S}$[three] $\cap$ ${}^{\leq S}$[three] $=$ [three] $\neq$ ${}^{\geq S}$[at least three] $\cap$ ${}^{\leq S}$[at least three] $=$ [at least three], because [three] $\neq$ [at least three].[11]

Temperature expressions Temperature scales are problematic for the standard Gricean picture. Intuitively, we conclude from *It is warm* that it is not hot, and from *It is cold* that it is not freezing. But what should the meanings be of these coarse-grained temperature expressions to predict these intuitions? The former inference is easy: we say that is warm if it is *at least*, say, 15 ^oC, and hot if it is, say, *at least* 25 ^oC. Thus, the proposition claiming that it is hot entails the proposition that it is warm, and we can by standard Gricean reasoning account for the first intuition. Moreover, by assuming that these expressions have at an 'at least'-meaning, we can immediately account for the phenomenon of *scale reversal*: the fact that from the scale ⟨boiling, hot, warm, ...⟩ we can derive ⟨..., not warm, not hot, not boiling⟩. This follows immediately by contraposition: if a entails b, by contraposition we immediately predict that $\neg b$ entails $\neg a$. Thus, from the assertion *It is not hot* we can derive by scalar implication that the stronger *It is not warm* is false, and thus that it is is warm but not hot, i.e., somewhere between 15 and 25 ^oC.

Unfortunately, to account for the intuition that from *It is cold* we conclude that it is not freezing, an 'at least' reading of coarse-grained temperature expressions gives exactly the wrong reading. Horn (1972) observed the problem and proposes to take the following two scales to be basic: ⟨boiling, hot, warm, lukewarm⟩ and ⟨freezing, cold, cool, tepid⟩. Although it is natural to assume that we do have these two scales, we would like to see how these scales are related to the meanings of the

[9] I owe his simple but still rather critical observation to Schulz (2001).

[10] For simplicity I leave out the rest of the sentence.

[11] Of course, the inference that ${}^{\geq S}$[three] $\cap$ ${}^{\leq S}$[three] $=$ [three] is based on our assumption that [three] $\in S$. From Merin (1999b) it is not very clear whether we are allowed to make this assumption, because although S plays a crucial role in Merin's formal analysis, he is normally not very explicit about how it looks like.

coarse-grained temperature expressions and on the basis of what principle the scales should be defined. Ducrot proposes an ordering in terms of argumentative value, and Merin (1999b, 2003) proposes a formal implementation. The reason why we have two scales is due to the fact that some (basic propositions based on) temperature expressions have a positive value with respect to some suitable chosen h (desiring for a high temperature), while others have a negative value (or positive w.r.t. the goal $\neg h$ for a low temperature). Moreover (though this is clearer in Ducrot's work than it is in Merin's), the use of a particular temperature expression, 'warm' versus 'cold', indicates what the speaker's preferences are, h versus $\neg h$. But how can we derive the two scales in terms of the meanings of the expressions? In analogy with his above described analysis of numerals, Merin crucially assumes that the meanings of the set of coarse-grained temperature expressions *partitions* the state space. Thus, 'warm', for instance, doesn't mean that it is *at least* 15 ^oC, but rather that the temperature is *between* 15 and 25 ^oC, leaving temperatures above 25 ^oC to be 'hot'. Although 'it is warm' is now taken to mean that the temperature is *between* 15 and 25 ^oC, what is asserted is still taken to mean that it is above 15 ^oC, because this is the upper cone of its semantic meaning. Still, if we also take the expected concession into account, we receive the desired reading: $^{\geq}$[It is warm] $\cap$ $^{\leq}$[It is warm] = [It is warm] $= \{t \in T|$ the temperature in t is between 15 and 25 $^oC\}$.

If scales are not defined in terms of entailment, the standard explanation for scale reversal can no longer be used. How then to account for it? At first, this seems rather straightforward, for according to the relevance function used by Merin, it holds that a is *more* relevant than b with respect to h if and only if a is *less* relevant than b with respect to $\neg h$. Unfortunately, given the 'exactly'-readings assumed by Merin, this doesn't really help him. Instead, Merin (1999b) proves and makes use of the following fact:

Fact 3 *If a and b are cells of a partition of the possibility space into propositions and $WA(h,a) > WA(h,b) > 0$, then $WA(\neg h, ^{<S} b) > WA(\neg h, ^{<S} a) > 0$ (and $WA(h, ^{\geq S} a) \geq WA(h, ^{\geq S} b) > WA(h,b)$).*

What we want to explain is that if a is 'better' than b, 'not b' is 'better' than 'not a'. Think now of 'not a' as as expression denoting $\overline{^{\geq S}[a]} = \ ^{<S}[a]$, and that it is made with respect to the opposite 'goal' as 'a' would be. Then it indeed follows that 'not b' is higher on the relevant scale than 'not a'.

Negation According to the standard Gricean picture, numerals, temperature expressions, and other phrases have semantically an 'an least'-meaning, and negation is a complement operator. This correctly predicts that *It is not warm* means that it is not warm and certainly not hot. Unfortunately, as observed by many scholars, we can also make statements like *It is not warm, it is hot!* This, of course, is problematic for the Gricean picture. The dilemma seems to be that we either have to give up the view that expressions have semantically an 'at least'-meaning, or that negation always denotes the complement operator. Thinking at the former as the heart and soul of Gricean pragmatics, Horn (1989) proposes to give up the latter view: negation does not always denote the complement operator, but can sometimes be used metalinguistically. In the problematic example, negation is used to claim that what was said previously was inappropriate.

Merin (2003) proposes another solution, making use of his upward and downward cones. Also he distinguishes two uses of negation, but hangs to the view that it always denotes a complement operation. According to him, the standard negation of ϕ is taken to be the complement of the upward cone (claim) of ϕ, $\overline{^{\geq S}\phi} = \ ^{<S}\phi$, while what Horn calls a meta-linguistic negation is proposed to be the complement of the downward cone (concession) of ϕ, i.e., $\overline{^{\leq S}\phi} = \ ^{>S}\phi$. Thus, negation is always taken to have the same meaning, but it gives rise to different results if applied to different propositions: a sentence's upward cone or its downward cone. In this way he can account for the appropriateness of both *It is not hot, but it is warm* (with standard negation) and *It is not warm, it is hot!* (with marked negation).

I believe that there are some problems with this view, however. First, it is not clear anymore what the second conjunct of the last example contributes to the meaning of the sentence: for $\overline{{}^{\leq S}[\text{It is warm}]} = {}^{>S}[\text{it is warm}]$, which, presumably, means that it is (at least) hot. A second and related observation is that the analysis becomes problematic when combined with Merin's own analysis of 'but'. Notice that *It is not warm, it is hot!* can naturally be reformulated as *It is not warm, but hot!*. But if the former conjunct already means that it is (at least) hot, in what sense could the second conjunct then be rhetorically opposite to the first?

Disjunction We saw above that in Merin's (1999b, 2003) analysis of conversational implicatures of numerals and temperature expressions it is assumed that these have a semantic 'exactly'-interpretation and because they are taken to be elements of the set S of relevant propositions in context that presumably partitions the state space, the intersection of upward and downward cones doesn't have any effect, and just produces the semantic meaning we started out with. More interestingly are cases where the semantic meaning of an expression is not taken to be an element of S.[12] Merin (1999a) discusses disjunction as such an example. In accordance with standard semantics, he takes the semantic meaning of a disjunctive sentence like 'a or b' to be that of inclusive disjunction. The set S is now explicitly claimed to be the following set of mutually incompatible propositions: $\{a \wedge b, a \wedge \neg b, \neg a \wedge b, \neg a \wedge \neg b\}$. Obviously, $a \vee b$ is no element of this set. To derive the exclusive interpretation of a disjunctive sentence, he proposes the upward cone of $a \vee b$ now to be the union of the first three propositions in S, while the downward cone is the union of the last three. Obviously, the net meaning, i.e., the intersection, contains only states in which exactly one of a and b is true, which is the desired exclusive reading. Unfortunately, Merin does not make clear why S should be as he proposes, nor why the upward cone, for instance, should be the union of $\{a \wedge b, a \wedge \neg b, \neg a \wedge b\}$. In particular, no mention is made of relevance to define this cone at all. Why should this be the default upward cone in the case of incomplete information, as suggested by Merin? Thus, while for numerals and temperature expressions almost nothing of interest happens when we go from semantic meaning to conversational implicature, now it is left completely unclear *why* the interesting inference to the exclusive interpretation is allowed.

Particularized scalar implicatures Now consider the particularized scalar implicature due to Mary's answer at her job interview to the question whether she speaks French by saying that her husband does. Obviously, the goal proposition, h, now is that Mary gets the job. Naturally, the proposition a = [Mary speaks French] has a higher relevance than the proposition b = [Mary's husband speaks French]. The net meaning of Mary's actual answer is claimed to be ${}^{\geq S}b = \bigcup\{s \in S | WA(h,s) \geq WA(h,b)\} \ \cap \ \bigcup\{s \in S | WA(h,s) \leq WA(h,b)\} = {}^{\leq S}b$. Now suppose that $b \in S$. Then, as always, $b \subseteq {}^{\geq S}b \cap {}^{\leq S}b$, and thus nothing is gained (though perhaps something is lost). Notice that if $b \in S$, the net meaning of b can only rule out that Mary herself speaks French, if this is already ruled out by the semantic meaning of b. So, if Merin assumed that $b \in S$, for the desired inference to go through, he also had to assume that the semantic meaning of b is something like 'Mary's husband speaks French and nobody else'. Given that Merin (1999b) claims that a and b are (presumably) logical independent, this cannot be what he had in mind. Perhaps Merin assumed with the rest of us that b has semantically speaking an 'at least'-meaning

[12]Strictly speaking, according to Merin's (1999b) definition 8 (our definition 3), upward and downward cones are not defined for propositions that are not elements of the set of relevant propositionns S.

saying that Mary's husband speaks French and perhaps others do as well. But then, of course, it has to be ruled out that $b \in S$. This could be done if we assume that S itself *partitions* the state space (and, as we have seen above, this is what he normally assumes, although he is almost never explicit about it). Presumably, this partition is induced by a question like *Who speaks French?* On this assumption it indeed follows that the elements of the partition compatible with $a =$ [Mary speaks French] are not compatible with the downward cone of b, and thus are ruled out correctly.

General scepticism Merin (1999a,b) gives an appealing analysis of the conventional implicatures associated with **adversary connectives** like *but* and other concessive-claim pairs. Indeed, I think that the work of Ducrot & Anscombre and of Merin strongly suggests that concession/claim pairs are more important in natural language use than Griceans normally admit (though see the remarkably honest Horn, 1991) and can be given a natural analysis by adopting an argumentative view on communication. Still, as it turns out, for these results the particular notion of relevance used by Merin is not required. Almost any quantitative notion of relevance has it that a is positively/negatively relevant to h if and only if $P(h \wedge a) > / < P(h) \times P(a)$,[13] and for all of these notions it is true that even if a entails b, $a \models b$, it can be that b is positively relevant to h, while a is negatively so. Something similar holds for Merin's derivation of the unexpectedness-appropriateness condition. Inspection of his proof shows that the above mentioned requirement for positive/negative relevance alone, together with the condition stating that 'a but b' is appropriate only if there is an h such that a is negatively relevant to h and b is positively relevant, suffices to derive the unexpectedness of b given a given the independence condition $[(a \perp b | \pm h)]_P$. The appropriateness condition doesn't have to be stated in terms of $WA(h, \cdot)$. For the default case $h = b$ this independence condition is not needed. Notice that if $h = b$, the sentence 'a but b' can be appropriate only if a is negatively relevant: $P(h \wedge a) < P(h) \times P(a)$. But this means that $P(b \wedge a) < P(b) \times P(a)$ and thus that $P(b/a) < P(b)$. Thus, Merin's treatment of 'but' does not depend on the particular notion of relevance he actually uses.

The above remarks show that Merin's analysis of adversary connectives does not rely on the particular relevance function chosen. They don't really diminish, however, the strength of his more general argumentative position. Merin (1999b) contrasts Grice's view of conversation as cooperative information exchange with his own argumentative view on conversation. Unfortunately, Merin's view – at least on its most straightforward reading – is in plain contradiction with general game theoretical results. As noted in section 2, Crawford & Sobel (1982) have shown that the amount of information that agents can transfer credibly in cheap talk games depends on the extent to which the preferences of the conversational agents are aligned. If the preferences are diametrically opposed, as proposed by Merin to be the default case, we would predict that no credible information can be communicated at all! Of course, we have also seen that communication is possible in real debates when it cannot be excluded that the information transferred might be falsified, and that communicating false 'information' is punished. The possibility of falsification and punishment make the communicative situations with seemingly non-aligned preferences ones where the preferences are more aligned after all, but then at a 'deeper' level. So I don't think that no sense can be made of Merin's bargaining view on communication at all, but he has to show us in what 'deeper' sense the preferences of the agents are still aligned.

But also if Merin can provide us with this missing link, his analysis of conversational implicatures I still find wanting. Although he typically can account for the phenomena, the hypothesis that conversational implicatures result from taking the intersection of the upward and downward cones is unconvincing.

[13] Next to the standard notion of relevance, $P(h/a) - P(h)$, and the one used by Merin, we also have $\frac{P(h/a)}{P(h)}$, $P(h \wedge a) - P(h) \times P(a)$, and $log\frac{P(h/a)}{P(h)}$.

First, Merin's proposal is counterintuitive in general: even if in actual conversation the goals and preferences of its participants are not always in complete alignment, we certainly do not always argue against each other. If I ask you who came to the party, it does not seem very natural to assume that we have dual preferences with respect to any relevant answer given (being sceptic doesn't mean having opposite preferences). Moreover, also in cooperative discourse situations 'net meaning' is stronger than what is explicitly asserted. If you answer my above question by saying that John and Mary came, I normally conclude that Sue didn't come, although this is not explicitly said. So, even if we would adopt Merin's proposal for argumentative discourse, we still need yet another approach towards implicatures in cooperative situations.

But even if we ignore this counterintuitive aspect of Merin's (1999b) proposal, his analysis of implicatures is still less than convincing. For numerals and course-grained temperature expressions the analysis already presupposes what should be explained,[14] and to arrive at the intuitively correct upward and downward cones we have to assume counterintuitive goal-propositions (the desire for many children and very high temperatures). For disjunction and particularized conversational implicatures his analysis crucially depends on the identity of the set S of relevant propositions, and he doesn't make it very clear *why* the chosen set should be chosen, nor why the upward and downward cones of a disjunctive sentence should be as he assumed. Moreover, for the analysis of all scalar implicatures, a typical Gricean requirement has to be made that the speaker be competent about the subject matter. It is not made clear why and how the inferences are cancelled or weakened in case this requirement is not met. Finally, although his uniform analysis of negation as complementation is an improvement on Horn's (1989) distinction between descriptive and metalinguistic negation, we will see soon that Horn's intuitions do not rule out a uniform analysis of negation as complementation either.

But Merin's analysis of scalar implicatures should not be ruled out just because it relies on a view of communcationn which is on one reading incompatible with general game theoretical results, nor because his treatment is, intuitively speaking, unconvincing. These 'higher-order' arguments can be used only if we can account for the same empirical phenomena in a theoretically more appealing way. In the next section I will set myself to this task by providing a, perhaps, more familiar and natural analysis of the implicatures discussed by Merin.

4 Implicatures and exhaustive interpretation

In section 2.2 we saw that it makes a lot of sense to assume that (truthful) speakers say as much as they can about for them desirable situations. In case the speaker is taken to be well-informed, we can conclude that what speakers do not say about desirable situations is, in fact, not true. We formulated a 'pragmatic' interpretation rule for sceptic hearers that have to 'decode' the message following this reasoning, to hypothesize what kind of situation the speaker is in. From now on, I will call this rule one of *exhaustive interpretation* formulated below. As in section 2.2, for this interpretation rule we can assume that '$t' < t$' if and only if the speaker prefers situation t to situation t'.

Definition 4 *Exhaustive interpretation (general)*

$$exh(A) = \{t \in [A] | \neg\exists t' \in [A] : t' < t\}$$

Now consider the example of scalar reasoning again that was a serious problem for standard Gricean analyses: the case where Mary answers at her job-interview the question whether she speaks French by saying that her husband does. Intuitively, this

[14] Except, of course, for 'at least'-expressions.

gives rise to the scalar implicature that the 'better' answer that Mary herself speaks French is false. As already suggested in section 3, this examples is only problematic for Gricean reasoning because that requires the relevant 'scales' to be defined in terms of entailment. But notice that if we assume the scale to be the preference order (between states) of the speaker, we can account for this example in terms of the rule of exhaustive interpretation. All we have to assume for this analysis to work is that the state where speaker Mary speaks French herself is more preferred to one where she does not. Thus, we can account for the particularized conversational implicature that Mary doesn't speak French in terms of the rule of exhaustive interpretation as already discussed in section 2.

Now I want to argue that the general principle of exhaustive interpretation defined above is not only relevant for cases where speaker and hearer have opposing preferences (to at least some degree), but that the principle is also perfectly applicable in ideal Gricean circumstances where the preferences of the agents are well-aligned.

In Gricean pragmatics, most conversational implicaturess – the scalar ones in particular – are due to the second submaxim of quantity, which requires a speaker to say as much about an issue as she knows to be true. Groenendijk & Stokhof's (1984) formulated an exhaustivity operator applied to (term) answers to questions that implements this principle in a natural way. As it turns out (van Benthem, 1989), this operator is virtually identical to McCarthy's (1980) rule of interpretation by predicate circumscription, which can be given a natural semantic characterization in terms of a minimal-, or preference-, model analysis. Assume that the (question)-predicate at issue is P ('who danced'), and that answer A ('John danced') is given. In that case, the exhaustive interpretation of A with respect to predicate P, or the circumscription of A with respect to P, will be the set of A-states where P has a minimal extension (John danced and nobody else):

Definition 5 *Exhaustive interpretation with respect to a predicate*

$$exh(A, P) = \{t \in [A] | \neg\exists t' \in [A] : t' <_P t\}$$

Thus the exhaustive interpretation of A contains all those states t that verify A, and for which no more minimal state t' exists that also verifies A. What still has to be specified is the ordering relation '$<_P$'. In standard circumscription it is assumed that $t' <_P t$ if the extension of P in t' is a subset of the extension of P in t, i.e. $P(t') \subset P(t)$.[15]

Obviously, definition 5 is formally very similar to definition 4: both interpretation rules are given a minimal model formalization. But the way these minimal models are defined seems different. In definition 4, the minimal models are the ones that are least preferred by the speaker. This is not the most straightforward interpretation of the ordering relation used in definition 5, but perhaps we can still provide one, and thus reduce the principle of exhaustive interpretation to the only rational interpretation rule in signaling games. For this derivation to go through, it has to be the case that $t' <_P t$ if and only if the speaker prefers t to t'. Why should that be so? Well, suppose that the preferences of the speaker are now well-aligned with those of the hearer, and that the latter (via her question 'Who has property P'?) wants to know the extension of P. But this means that if $t' <_P t$, the speaker can say in t' of less individuals that they have property P than in t, and thus can inform the hearer of less individuals with certainty that they have property P. Given that the hearer wants to know of as much as possible individuals whether they have property P, and that the making of negative statements is, if possible, to be avoided,[16] it follows that the speaker prefers

[15]It is also assumed that with respect to all other predicates P', $P'(t') = P'(t)$. This is important for the circumscription analysis of conditionals, for instance, but won't bother us in this paper.

[16]This, obviously, is the weak point of the argument. But I believe that this is a conventional feature of natural languages and is defensible for reasons of *efficiency*. See van Rooy (2004) for more discussion.

t to t' if $t' <_P t$.

The principle of exhaustive interpretation will be used to account for scalar implicatures. In definition 5 it crucially relies on a particular question predicate. More in accordance with, perhaps, more standard analyses of Gricean implicatures (e.g. in the work of Horn (1972, 1989), Gazdar (1979), and many others), we can also make exhaustification relative to a particular language in question. The standard analysis of exhaustification can be described alternatively relative to a particular language L as follows:

Definition 6 *Exhaustive interpretation with respect to a language*

$$exh(A, L) = \{t \in [A] | \neg\exists t' \in [A] : \{B \in L | t' \in [B]\} \subset \{B \in L | t \in [B]\}\}$$

Obviously, these definitions are equivalent if we take language L to be defined in terms of predicate P as follows: $L = \{P(a) | a \text{ is an individual constant}\}$, if we assume that every individual has a (unique) name.[17]

It is well-known that the above rule of exhaustive interpretation can account for many conversational implicatures, in particular for a number of *scalar* ones (see Groenendijk & Stokhof (1984), van Rooy (2003), van Rooy & Schulz (submitted), and Spector (manuscript)). But we want to know whether in terms of exhaustive interpretation we can account for those phenomena for which Merin (1999b) claims an argumentative view on language use is required. If we could do so, it shows that, perhaps, Merin's argumentative view on language use need not have such a wide domain of application as he explicitly assumes himself, i.e., that *all* language use is argumentative. We will take up the phenomena discussed in section 3.2 one by one. I would like to mention that almost none of the assumptions I make below are particularly new or surpising, though perhaps it has not been very clear how they can be used to describe the phenomena at hand.

Numerals As we saw in the previous section, although standard Gricean pragmatics can account for the (default) inference that John has exactly three children from the assertion *John has three children*, it has problems with sentences like *John has <u>at least</u> three children.* Indeed, on the assumption that 'three children' and 'at least three children' have the same semantic meaning, the standard analysis of exhaustive interpretation (Groenendijk & Stokhof, 1984) has problems to account for the fact that the exactly-interpretation is missing for the version modified by 'at least'. But then, like Merin (1999b, 2003), we can simply assume that *three* means something different than *at least three.* Indeed, let us just assume (temporarily) that [three] = 3. Moreover, let us assume that 'John has three children' is true in state t if 3 is an element of the extension of property λn[John has n children] in state t. In general, we assume that $P(3)$ is true in t iff $3 \in P(t)$. Of course, we want to assume that if John has 3 children, he has 2 children as well. This can be accounted for by assuming that property λn[John has n children] is monotone *downward entailing.* A property P is downward entailing with respect to numerals iff it holds for all natural numbers n and m and states t that if $n \in P(t)$ and $m < n$, then also $m \in P(t)$.[18] Thus, for downward entailing P, $P(3)$ entails $P(2)$, but the truth of $P(2)$ does not rule out that also $P(3)$ is true. In this sense numerals are still predicted to have an 'at least' reading

[17] In this case, the sentences in L are logically independent to one another. In fact, it has been repeatedly suggested that the set of alternatives should consists of propositions that are all mutually independent in this way. As we will see later, compatibility is a natural demand to make on the members of L, but non-entailment is not, or so I believe. Problematic for the latter constraint are numerals and temperature expressions, for instance, discussed shortly. I take *four* and *hot* to be alternatives of *three* and *warm*, respectively (and vice versa), although the former ones entail the latter ones (for numerals, with respect to monotone decreasing predicates).

[18] To guarantee monotonicity properties of predicates of a language, we have to make use of meaning postulates, i.e., constraints on the structures that can be used to model the language.

when used with a downward-entailing predicate. Obviously, we also have predicates that denote monotone upward entailing properties, and properties that don't behave monotonically at all. An example of the former kind of predicate is 'Can run the 100 meters in n seconds', while mathematical predicates like '$3 + n = 7$' are prime examples of expressions denoting non-monotonic properties.

This is all fine, but how does this analysis account for the fact that *John has <u>at least</u> three children* doesn't give rise to the implicature that he has no more than three children? To account for this, we have to slightly change the above analysis: we don't say anymore that 'three' means 3 and that P(three) is true in t iff $3 \in P(t)$, but rather with Kadmon (1987) that 'three' means $\{x|card(x) = 3\}$ and that P(three) is true in t iff $\exists x \in$ [three] : $x \subseteq P(t)$. Similarly, we say that 'at least three' means $\{x|card(x) \geq 3\}$ and that P(at least three) is true in t iff $\exists x \in$ [at least three] : $x \subseteq P(t)$. Van Rooy & Schulz (submitted) show that in combination with some standard assumptions in dynamic semantics, this analysis leads to the correct prediction that while the exhaustive reading of P(three) gives rise to an exactly reading, this is not so for P(at least three).[19] Because I cannot assume the readers of this journal to be familiar with dynamic semantics, I must leave the details of the analysis to the paper mentioned.

Temperature expressions Remember that we normally conclude from an assertion *It is warm* that it is not hot, while from *It is cold* we conclude that it is not freezing. The problem here – as already noted by Horn (1972) – is that on a standard 'at least'-interpretation of temperature expressions ('warm' means higher than or equal to 15 ^0C, and 'cold' means higher than or equal to 5 ^0C), only the first inference can be accounted for by standard Gricean reasoning. Merin (1999b) proposes that temperature expressions give rise to two scales, and that the division is based on which temperature statements are positively relevant with respect to two opposing 'goals' h and $\neg h$. Indeed, in this way the two desired scales ⟨boiling, hot, warm, lukewarm⟩ and ⟨freezing, cold, cool, tepid⟩ *can*, in principle, be derived. However, it is not made clear which particular propositions *should* be at issue to derive those orderings, nor why *only* these propositions *could* be relevant.

Though Merin (1999b) bases his analysis on Ducrot's general argumentative view on language, Ducrot himself thinks of this argumentative orientation somewhat differently from Merin. While Merin takes the argumentative orientation with respect to a *contextually given* proposition, Ducrot assumes that the argumentative orientation is *inherent* to the *language* itself, and accordingly is context independent. Thus, it is expressions themselves that have, according to Ducrot, already a certain argumentative orientation. To take one of Ducrot's favorite examples, whereas *little* has a negative argumentative orientation and can be used to argue for 'negative' conclusions, *a little* has a positive one. The same can be said for *warm* versus *cold*, they have *as part of their meaning* already an argumentative orientation. In logical terms this is normally expressed by saying that whereas the one has a monotone *increasing* meaning, the other is monotone *decreasing*. Suppose that also *warm* and *cold* have a conventional, or 'semanticized', argumentative orientation: the former 'goes to' high temperatures and the latter to low ones. But then we can state the meanings of the two expressions by [warm] $= \{n|n \geq 15^0C\}$ and [cold] $= \{n|n \leq 5^0C\}$. The scalar inferences can now be accounted for by standard exhaustive reasoning.

Let us assume that a statement like *It is warm/cold* is normally made if the

[19]True, on this analysis we have to assume that a quantified expression like 'at most three' should be defined in terms of negation as something like 'not (at least) four'. Our analysis crucially relies on a distinction between indefinite and quantificational noun phrases: in dynamic semantics, only the former immediately introduce discourse markers. Alternatively (as suggested to us by Cornelia Endriss), we could make the distinction differently and say that both 'At least/most three men came' should be seen as quantification sentences interpreted as $\exists x[x \in$ [at least/most three] & $x = (Men \cap Came)]$, with '=' instead of '$\subseteq$'. Now we don't have to decompose 'at most three'.

temperature is at issue. So, to account for the implicatures it seems natural to assume that the relevant language L consists of coarse-grained temperature expressions like *freezing*, *cold*, *cool*, *tepid*, *warm*, *hot*, and *boiling*. We will assume the following meanings of these expressions: [freezing] = $\{n|n \leq 0^0C\}$, [cold] = $\{n|n \leq 5^0C\}$, [cool] = $\{n|n < 10^0C\}$, [tepid] = $\{n|n \geq 10^0C\}$, [warm] = $\{n|n \geq 15^0C\}$, [hot] = $\{n|n \geq 25^0C\}$, and [boiling] = $\{n|n \geq 100^0C\}$. Notice that from these meanings we can derive the following entailment relations: boiling $\models$ hot $\models$ warm, and freezing $\models$ cold $\models$ cool, but that no inference relations exist between warm and cold, for instance. In fact, if two of the above expressions do not stand in an inference relation to one another, they are incompatible.

Applying now the exhaustivity operator with respect to a language,[20] and assuming that the language L contains of the coarse-grained temperature expressions mentioned above, we immediately can infer from *It is warm* that it is not hot, and thus that it is between 15 ^0C and 25 ^0C. Similarly, we can infer from *It is cold* that it is not freezing, and thus that it is between 5 ^0C and 0 ^0C.

Now consider what we predict for negative statements like *It is not hot*. This, of course, depends on what we take to be the relevant language. Suppose that we take the same language as we had before. Then we falsely predict that from *It is not hot* we conclude that it is freezing! The assumption that for negative statements we switch to the following language: $\bar{L} \stackrel{\text{def}}{=} \{\bar{a}|a \in L\}$ (where $\bar{a}$ denotes the complement of a) also doesn't make a lot of sense. It would lead to the false prediction that it is somewhere between 5 ^0C and 15 ^0C. What we like to predict is that it is warm, i.e., somewhere between 15 ^0C and 25 ^0C, and we make this correct prediction if we assume that (i) the relevant language is now the following subset L' of L: {boiling, hot, warm, lukewarm} and (ii) that for a negative sentence we now take $\bar{L'} = \{\bar{a}|a \in L'\}$ of this relevant subset. We can derive the particular subset L' by assuming the following constraint on a language with respect to which we have to apply exhaustification: all elements of L' have to be mutually compatible. This constraint gives us the correct prediction for 'negative' temperature statements, and it doesn't do any harm for their 'positive' counterparts, so it seems only natural to use this constraint as a strong requirement for all cases.

In conclusion: we don't have to associate an argumentative orientation with temperature expressions to account for the desired conversational inferences. It is enough to assume that some expressions have a 'at least' and others an 'at most' meaning, and that temperature statements are interpreted exhaustively with respect to a suitable language L.

Negation Recall the distinction between *It is not hot, but it is warm* (with standard negation) and *It is not warm, it is hot!* (with marked negation). On our analysis of 'hot' and 'warm' it follows that while in the former case the negation can have a straightforward descriptive reading, in the latter case it cannot. Horn (1989) proposes to make a distinction between normal and metalinguistic negation. But on this analysis it isn't clear anymore that negation just denotes complementation. Merin (2003) proposes to account for the distinction by assuming that negation is a complement operator that applies either to the upward cone of a sentence, as in the former case, or to its downward cone. But Van der Sandt (1991) showed already that we don't have to take upward and downward cones into account to describe the distinction between the two uses of negation without giving up the idea that negation denotes complementation. For our case at hand, we can just assume that in one kind of use, negation takes the complement of the standard semantic meaning, while in the other use it denotes the complement of the exhaustive meaning. Thus, in the sentences above, it denotes $\overline{[hot]} = \{n|n < 25^0C\}$ in the first example, and

[20] Of course, doing things in terms of exhaustivity with respect to a predicate would be equivalent.

$\overline{[exh(warm, L)]} = \{n | n < 15^0C \text{ or } n \geq 25^0C\}$ in the second. Notice that in contrast to Merin's analysis, we don't have any problem to explain why the second conjunct *it is hot!* can still be claimed informatively in the second example.

Disjunction Let us now consider the scalar implicature associated with disjunctive sentences that they should be read exclusively. How can we derive the exclusive reading of a disjunctive sentence by applying an exhaustivity operator? Obviously, we immediately receive the correct results if we interpret a sentence of the form '$P(a) \vee P(b)$' exhaustively with respect to predicate P. Notice that this gives the correct result also for '$P(a) \vee P(b) \vee P(c)$'. However, or at least so is believed by most authors, the inference to the exclusive reading seems to be relatively context-independent, i.e., not dependent on whether a particular predicate is at issue. Moreover, in this way we can't derive the exclusive reading for a disjunctive sentence like '$P(a) \vee \neg P(b)$'. A natural proposal to solve both problems at once is to take the relevant language L now to be just the set of disjuncts occurring in the disjunctive sentence (e.g. Spector, 2003). For a sentence like $a \vee b$, this means that $L = \{a, b\}$, while for $a \vee \neg b$ it is $\{a, \neg b\}$. When we now apply exhaustification to a disjunctive sentence A with taking L to be the set of disjuncts containing in A, we immediately receive the desired exclusive reading.

Particularized scalar implicatures Earlier in this section we accounted for the implicature arising from Mary's answer at her job-interview in terms of the speaker's preference relation among states. Merin (1999b) proposed that this example should be accounted for in terms of his relevance relation $WA(\cdot, \cdot)$, while we have shown above how the use of a language, or a set of alternatives, can be useful for the analysis of implicatures. Here I want to suggest that we can also account for Mary's answer by taking relevance and a particular language into account.

In definitions 5 and 6 we defined our exhaustivity operator (implicitly or explicitly) in terms of the standard notion of entailment. However, we can generalize this operator by taking a notion of relevance into account. In general we can use the following exhaustivity operator, which depends not only on a language L, but also a goal proposition h and a relevance function V:[21]

Definition 7 *Exhaustive interpretation with respect to relevance*

$$exh(A, L, h) = \{t \in [A] | \neg \exists t' \in [A] : t' <_h^L t\}$$

In this definition we assume the following ordering relation between states:

$$t' <_h^L t \quad \text{iff} \quad V(h, \bigcap\{B \in L | t' \in [B]\}) < V(h, \bigcap\{B \in L | t \in [B]\})$$

Before we return to Mary's job interview, let us first assure ourselves that this new exhaustivity operator is really a generalization of our earlier ones. But this follows immediately once we take h normally to be $\bigcap L$.[22] Thus, if we use standard predicate circumscription with respect to predicate P, $\bigcap L$ denotes the proposition $[\forall x P(x)]$; for a sentence like *It is warm*, this intersection will denote that it is very hot; and for a disjunctive sentence it denotes the proposition that all disjuncts are true. Consider now the set $\bigcap\{B \in L | t \models B\}$ in case predicate P is at issue. It denotes the proposition that *at least* all individuals that have property P in state t actually have property P. That is, it denotes the following proposition: $\lambda i[P(t) \subseteq P(i)]$. Making use of exactly this proposition, van Rooy & Schulz (submitted) show that the ordering relation

[21]This could, but need not, be the one used by Merin (1999a,b).

[22]Notice that this results in the empty set if the elements of L are not mutually compatible, something that we actually ruled out. But even then this goal proposition need not be natural, but I just want to show that we could do things in terms of goals as well.

$t' < t$ between states mirrors an entailment relation exactly when all information is relevant. In particular, this shows that in our special case $V(h, \bigcap\{B \in L | t \in [B]\}) < V(h, \bigcap\{B \in L | t \in [B]\})$ if and only if $\{B \in L | t \in [B]\} \subset \{B \in L | t \in [B]\}$. Thus, in the special case under discussion, definition 7 reduces to definitions 5 and 6.

However, by making exhaustive interpretation dependent on relevance, we can account for more phenomena than we could until now. As shown in detail in Van Rooy & Schulz (submitted), now we can also account for some phenomena which are standardly taken to be problematic for the principle of exhaustive interpretation. In particular, we can now account for the inference that from Mary's answer in her job interview we conclude that she doesn't speak French herself. The reason is, obviously, that the goal proposition that Mary speaks French is also an element of L and more valuable than the proposition saying that her husband does.

In this paper I have argued against Merin's analysis of scalar implicatures in terms of his upward and downward cones. But, by taking relevance into account, for many examples an analysis in terms of exhaustive interpretation is in fact very similar to what Merin proposed. By a clever – though sometimes unnatural – choice of goal proposition h, it is in most (if not all) cases possible that what we take to be the semantic meaning of an expression with a monotone increasing meaning is the same as what he takes to be the upward cone of the proposition expressed by a phrase with a (semantically) non-monotonic meaning. Let denote this proposition by A. In those cases, Merin's intersection of this resulting upward cone with a proposition's downward cone will be the same as our exhaustive interpretation of A.

5 Conclusion

It is commonly assumed that the possibility of linguistic communication depends crucially on mutual cooperation between members in a group. Lewis (1969) assumed alignment of preferences to account for linguistic conventions and we have seen that later work in economics proved that such an alignment of preferences is indeed required for reliable communication to be possible. But this proof is based on the assumption that (the preferences of the agents are common knowledge and that) the informational content of a message is not verifiable. Perhaps, if we drop this limitation, we can think of actual communication as a (to a large degree) non-cooperative affair after all.

In this paper I have contrasted the standard cooperative view on communication with an argumentative one. According to the former view, we communicate information that is good for all participants of a conversation, while according the latter, we communicate always to argue for a particular hypothesis and do this always against an opponent. In this paper I have put some doubts on the universal applicability of either view. Though an argumentative perspective was conceded to be useful for the analysis of (at least) adversary connectives, adopting this perspective for the analysis of all scalar implicatures (or at least in the way proposed by Merin) was argued to be unconvincing. I have argued that a more natural explanation is possible when we assume speakers to say as much as they can, and hearers to interpret exhaustively. Although this assumption seems rather standard, it does not require perfect alignment of preferences as normally presupposed in Gricean pragmatics. Thus, even if neither the Gricean cooperative view on language use, nor the alternative argumentative view has universal applicability, this doesn't mean that conversational implicatures cannot still be accounted for by means of a general rule of interpretation. Obviously, I take the principle of exhaustive interpretation to be such a general rule.

References

[1] Anscombre, J.C. & O. Ducrot (1983), *L'Argumentation dans la Langue*, Brussels:

Mardaga.

[2] Benthem, J. van (1989), 'Semantic parallels in natural language and computation', In H. D. Ebbinghuas et al. (eds.) *Logic Colloquium '87*, Elsevier Science Publishers, Amsterdam, pp. 331-375.

[3] Crawford, V. and J. Sobel, (1982), 'Strategic information transmission', *Econometrica*, **50**: 1431-1451.

[4] Ducrot, O. (1973), *La preuve et le dire*, Paris: Mame.

[5] Fauconnier, G. (1975), 'Pragmatic scales and logical structure', *Linguistic Inquiry*, **6**: 353-375.

[6] Frege, G. (1918), 'Der Gedanke: eine logische Untersuchung', *Beitrage zur Philosophie des deutschen Idealismus*, **1**: 58-77.

[7] Gazdar, G. (1979), *Pragmatics*, Academic Press, London.

[8] Gibson, R. (1992), *A Primer in Game Theory*, Harvester Wheatsheaf, Hertfordshire.

[9] Glazer, J. and A. Rubinstein (2001), 'Debates and decisions: On a rationale of argumentation rules', *Games and Economic Behavior*, **36**: 158-173.

[10] Good, I.J. (1950), *Probability and the Weighing of Evidence*, Griffin, London.

[11] Grice, H.P. (1967), 'Logic and conversation', *William James Lectures*, Harvard University, reprinted in *Studies in the Way of Words*, 1989, Harvard University Press, Cambridge, Massachusetts.

[12] Groenendijk, J. and M. Stokhof (1984), *Studies in the Semantics of Questions and the Pragmatics of Answers*, Ph.D. thesis, University of Amsterdam.

[13] Hirschberg, J. (1985), *A theory of scalar implicatures*, PhD thesis, University of Pennsylvania.

[14] Horn, L. (1972), *The semantics of logical operators in English*, PhD Thesis, Yale University.

[15] Horn, L. (1989), *A Natural History of Negation*, University of Chicago Press, Chicago.

[16] Horn, L. (1991), 'Given as new: When redundant affirmation isn't', *Journal of Pragmatics*, **15**: 313-336.

[17] Kadmon, N. (1987), *On Unique and Non-Unique Reference and Asymmetric Quantification*, Ph.D. thesis, University of Massachusetts, Amherst.

[18] Lewis, D. (1969), *Convention. A Philosophical Study*, Harvard University Press, Cambridge, Massachusetts.

[19] Lipman, B. (2002), 'Language and Economics', In M. Basili et al. (eds.), *Cognitive Processes and Rationality in Economics.*

[20] Lipman, B. and D. Seppi (1995), 'Robust inference in communication games with partial provability', *Journal of Economic Theory*, **66**: 370-405.

[21] Maynard Smith, J. & E. Szathmary, (1999), *The Major Transitions in Evolution*, Oxford University Press.

[22] McCarthy, J. (1980), 'Circumscription - a form of non-monotonic reasoning', *Artificial Intelligence*, **13**: 27-39.

[23] Merin, A. (1995), 'Why not Kim Basinger?', In: H. Kamp & B. Partee (eds.), *Context-Dependence in the Analysis of Linguistic Meaning: Proceedings of the Workshop in Prague/Bad Teinach*, Stuttgart University.

[24] Merin, A. (1997), 'If all our arguments had to be conclusive, there would be few of them', *Arbeitspapiere des Sonderforschungsbereichs 340*, Nr. 101, Stuttgart.

[25] Merin, A. (1999a), *Die Relevance der Relevance: Fallstudie zur formalen Semantik der englischen Konjuktion* but, Habilitationschrift, University Stuttgart.

[26] Merin, A. (1999b), 'Information, relevance, and social decisionmaking', In: L. Moss, J. Ginzburg, M. de Rijke (eds.), *Logic, Language, and Computation. Vol. 2*, Stanford.

[27] Merin, A. (2003), 'Replacing 'Horn scales' by act-based relevance orderings to keep negation and numerals meaningful', In: reader for course *Relevance and Decision-Theoretic Semantics*, given by A. Merin at ESSLLI summerschool in Vienna, 2003.

[28] Parikh, P. (2001), *The use of Language*, CSLI Publications, Stanford, California.

[29] Reichenbach (1956), *The Direction of Time*, University of Chicago Press, Berkeley.

[30] Rooy, R. (2003), 'Relevance and bi-directional OT', In: R. Blutner and H. Zeevat (eds.), *Pragmatics in Optimality Theory*, Palgrave Macmillan, pp. 173-210.

[31] Rooy, R. (2003), 'Being polite is a handicap', In: Tennenholtz et al. (eds.), *Proceedings of TARK 9: Theoretical Aspects of Rationality and Knowledge*. Bloomington.

[32] Rooy, R. (2004), 'Evolution of conventional meaning and conversational principles', *Synthese*, in print.

[33] Rooy, R. van & K. Schulz (submitted), 'Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation'.

[34] Rubinstein, A. (2000), *Economics and Language*, Cambridge University Press, Cambridge.

[35] Sandt, R. van der (1991), 'Denial', *CLS 27, Parasession on Negation.* Chicago Linguistic Society, Chicago, pp. 331-334.

[36] Schulz, K. (2001), *Relevance and 'Quantity' Implicaturen*, Diplomarbeit, University of Stuttgart.

[37] Spector, B., (manuscript), 'Scalar implicatures: local or global?', Handout Diptalk at the University of Amsterdam.