

## Chapter 16

### Accommodation, Meaning, and Implicature: Interdisciplinary Foundations for Pragmatics

*Richmond H. Thomason*

---

#### 1 Introduction

An interest in pragmatics generally arises through concern with the problems of some related field. In my case the related field was semantics. My interest in linguistic meaning originated in philosophical logic; the goal was to develop a logically based semantics in the spirit of Carnap 1947. This methodology was particularly narrow in relation to linguistic evidence. Philosophers in general, and philosophical logicians in particular, tend to concentrate on a few key examples that illustrate deep theoretical issues (though many of these examples seem fairly artificial, from a linguistic standpoint). The philosopher's game is to grapple in the theoretical arena with these few examples. This can lead to deep theories; but considered as theories of language the products of this methodology tend to have a certain artificiality and instability.<sup>1</sup>

Involvement with Montague Grammar acquainted me with semantics as practiced by linguists, and with much more demanding standards of how theories should be related to language. Thinking about examples was not new to me. But the breadth of contact with linguistic evidence, and the goal of putting theories in a systematic relation with this evidence, did seem new.

The tradition that developed out of Montague's work is now a small, but relatively healthy subfield of linguistics, that has generated a considerable literature of its own. Despite the record of progress on specific research topics that made these developments possible, this work hasn't created the sort of secure basis that would enable you, for instance, to write an

Ultimately, this article is a descendent of an unpublished manuscript, called "Semantics, pragmatics, conversation and presupposition," which I presented at the University of Texas in 1973, and circulated privately. Although I revised that paper several times and presented it elsewhere, it remained unpublished because I felt that the theoretical ideas were insufficiently developed. The approach that I take here owes much to the subsequent work of others. I try to acknowledge these debts as they come due in the course of the exposition; in particular, though, I'm indebted to Andrew McCafferty for general conversations on the topic, and for specific comments on drafts of this paper. This version of the paper was written in 1987.

introductory textbook with a sense of confidence in the stability of the material. Moreover, quite basic theoretical issues regarding commonplace linguistic phenomena remain largely unresolved: I am thinking of the interpretation of constructions like the conditional, generic plural, the definite article, and aspect.

In a way, it isn't surprising that attempting to apply these more demanding linguistic standards of evidence to semantic theories wouldn't meet with unqualified success. If this were a straightforward matter, the philosophy of language wouldn't have split earlier in this century into a "logical" school and an "ordinary language" school. During this time there were many clever people thinking about language, and someone would probably have noticed that the division was unnecessary.

Of course, philosophical logicians do recognize the need to relate their theories to evidence having to do with reasoning and language. This accounts for much of the interest in "truth conditions," since one way to refute a theory of validity is to show that some of the inferences it recommends are in fact invalid—that the premisses can be true while the conclusion is false.

But truth conditions are not transparent to observation. Our direct intuitions about the truth and falsity of sentences hold up in circumstances that are too circumscribed to test theories as we would like to test them. When we know all the relevant propositions, and are given enough information about a context, and we are given a sentence that would be appropriate to assert in the context, we can generally agree about whether the sentence is true. But in testing logical theories, the crucial cases are often sentences that would be inappropriate if uttered. And we seem to lack the intuitions that are needed to test validity. Though this has been a major and recurrent difficulty in logic, philosophical logicians have been slow to meet the methodological challenge.

To illustrate the historical point, I'll briefly consider three case studies: Strawson's criticism of Russell's theory of definite descriptions, the motivation of relevance logic, and Grice's causal theory of perception. Readers who are not interested in these details can pass over the next section.

## 2 *Philosophy of Language and Evidence for Truth*

### 2.1 *Definite Descriptions*

It is difficult to untangle the theoretical issues raised by Strawson (1950). I want to focus only on the methodological issue of what linguistic evidence Strawson is able to present for his conclusion that 'The king of France is wise' is neither true nor false. What this comes down to is an invitation to imagine that, with all appearances of seriousness, someone utters this

sentence to you.<sup>2</sup> Strawson performs this experiment, and then makes the following points:

1. You wouldn't respond "That's untrue";
2. If you were asked to say whether you agreed or disagreed with what was said you would say (perhaps with some hesitation) that you didn't do either;
3. You would take someone's assertion of 'The king of France is not wise' to be evidence that he believed that there is a king of France.

Russell, of course, was skeptical about what could be learned about logic from usage. However, in his reply to Strawson he does make one point bearing on the linguistic evidence.

... though I have no wish to support the claim of common usage, I do not think that [Strawson] can claim it either. Suppose, for example, that in some country there was a law that no person could hold public office if he considered it false that the Ruler of the Universe is wise. I think an avowed atheist who took advantage of Mr. Strawson's doctrine to say that he did not hold this proposition false would be regarded as a somewhat shifty character. (Russell 1970, 389)

Neither of the parties in this dispute explicitly confronts the thinness of Strawson's approach to matters of linguistic truth and falsity. The issue of *dissembling* that is raised by Russell's example illustrates the difficulty. Whether or not Russell's atheist is dissembling, we do recognize that dissembling occurs: there are cases in which you can mislead an audience by uttering sentences that nevertheless are true. And once we see this, we can immediately think of many other legitimate ways in which a true utterance can be criticized. Thus, in merely citing our reluctance to assert the sentence, Strawson has made an incomplete case that 'The king of France is not wise' is untrue. He must show that this reluctance is due to the sentence's untruth rather than to some other factor for which we could be criticized in asserting the sentence.

What is needed is some way of testing anomalous utterances for truth and falsity. Direct reflection is obviously of little use, since Russell and Strawson seem to have honestly had opposite intuitions in this case.

At this point, purely philosophical methods seem to have reached an impasse; I myself do not think that the matter can be taken further without an entirely different approach to the evidence.

### 2.2 *Relevance and Conditionals*

The theory of the conditional that is developed in Anderson and Belnap 1975 subscribes to the generalization that a conditional is untrue if the antecedent is not relevant to the consequent. The book is largely devoted

to technical developments, and like many contributions to philosophical logic does not attempt to give a close examination of the linguistic evidence. But the relevance claim is illustrated by the following example: a mathematician writes a paper in which he makes a conjecture about Banach spaces, and then goes on to write

In addition to its intrinsic interest, this conjecture has connections with other parts of mathematics that might not occur to the reader. For example, if the conjecture is true, then the first order functional calculus is complete . . . . (Anderson and Belnap 1975, 17)

Anderson and Belnap's point is that the editor would be right in asking the author to remove this claim from the paper; perhaps, as they say, the editor may even remark that the journal requires the antecedent of an asserted conditional to be relevant to its conclusion.

But a moment's thought makes it clear that there are many true mathematical claims that an editor would not allow in a publishable mathematics paper. Trivial claims are certainly among these. A theory of the conditional that makes the author's claim trivial would explain the editor's reaction just as well as one that makes it untrue.

Again, if we directly confront the issue of whether irrelevant conditionals are thereby untrue, we are forced to judge the truth conditions of utterances that in general will be inappropriate. And the issue is just as unclear.

For instance, "monkey's uncle" conditionals provide some evidence that irrelevant conditionals are sometimes treated as true. But as far as I know, there has been no definitive or really systematic study of the evidence on irrelevant conditionals. Relevance logic remains unsupported (and also unrefuted) by the linguistic evidence.

### 2.3 *The Causal Theory of Perception*

In Grice 1961, Grice makes methodological points very like the ones that I have made, but in the context of ordinary language philosophy rather than logic. He assumes that linguistic usage is an important test for philosophical theories, and addresses in particular a theory postulating sense-impressions that are caused by material objects in the course of perceptual occurrences. It is crucial to such a theory that statements such as "That looks red to me" are true even in the normal case when, without any suggestion of doubt or dispute, one is face-to-face with an obviously red object. However, statements of this kind are obviously anomalous.

This anomaly was one of the chief weapons invoked by J. L. Austin in his demolition of sense-datum theory.<sup>3</sup> Like Strawson, Austin makes much of the anomaly without directly recognizing truth as an independent issue, and confronting this issue.

For our purposes, the important parts of Grice 1961 are sections 3 and 4 pages 126–146. Grice prefaces these two sections by saying that for a long

time he saw no way to use linguistic phenomena to resolve the question of whether these anomalous statements are neither true nor false, or are true but misleading—but that now he believes that this issue can be settled in favor of the latter position. Section 3 consists of a statement of the theory of implicature, including linguistic tests for distinguishing conventional from conversational implicatures. The main point of section 4 is to apply linguistic tests to 'That looks red to me', arguing from the results of these tests that the implicature is conversational.

Here, we have a case in which a philosopher of language does squarely face the issue of truth. And there has been progress; in fact, Grice has proposed his theory of implicature in these sections.<sup>4</sup> But the application of the theory to the philosophical issue is less satisfactory than one would like. In the last analysis, Grice's argument for the truth of statements like "That looks red to me" when 'that' refers to a ripe tomato depends on a highly indirect argument from simplicity; there is a general theory of conversational effects on utterances that predicts that these statements will be anomalous, even if they have the truth conditions that the sense-datum theorist requires. This theory is needed anyway for other purposes, and is explanatory. An alternative theory of truth conditions that invokes truth-value gaps would be more complex than the sense-datum theorist's. Therefore, on the whole, it would be better to say that these statements are true, rather than neither true nor false.

The argument depends on the simplicity and generality of a combination of a two-valued semantic theory with a pragmatic theory of implicature that predicts anomaly. The problem, however, is that simplicity and generality are not the only desirable qualities in a theory; rigor and precision of explanation count for something as well. And though the theory that Grice offers is very plausible, it is too loose-knit to support a good argument from simplicity and generality. Though they are unsatisfactory in many ways, rules for manipulating truth-value gaps can at least be made explicit and precise.<sup>5</sup> The most pleasant solution to this difficulty would be to preserve the generality and plausibility of Grice's approach, while achieving a more explicit theory, and a much finer texture of explanation. I will urge in this paper that by putting together developments in computer science, linguistics, and philosophy we have some hope of making incremental progress toward this goal.

### 3 *Prospects for Pragmatics*

Whether one is a philosopher or a linguist, the process of trying to justice to evidence in cases such as those that I have illustrated quickly convinces one that simple theories of truth and falsity, even if they are made context relative, fail to do justice to the phenomena.<sup>6</sup>

*Presupposition* is one of the most obvious illustrations of the point. In fact, the use of a nonsemantic strategy of accounting for presupposition goes back about as far as it can, to Frege.

If one therefore asserts 'Kepler died in misery', there is a presupposition that the name 'Kepler' designates something; but it does not follow that the sense of the sentence 'Kepler died in misery' contains the thought that the name 'Kepler' designates something. If this were the case the negation would have to run not

Kepler did not die in misery

but

Kepler did not die in misery, or the name 'Kepler' has no reference. (Frege 1960, 69)

But if pragmatics rather than semantics is responsible for presupposition, we have to say exactly how it is going to carry out its responsibilities. If linguistic evidence must be confronted by a combined theory that incorporates both semantic and pragmatic elements, we need to have some sense of how explanatory duties will be allocated. Briefly, I have in mind the following picture.<sup>7</sup>

It is semantics that has to do with the interpretation of phrases relative to a context of utterance. Pragmatics should deal with the *use* of interpreted phrases: with what acts a speaker performs by using a phrase in a certain context. Among these acts are things like saying, implicating, insinuating, and presupposing<sup>8</sup>—thus, the message that a speaker *gets across* by saying something in a given context also belongs to pragmatics. And in fact, this is the part of pragmatics that I want to concentrate on here.

If, as Grice feels, the source of such phenomena lies in *principles of rational conversation*, it should be possible to single out certain important types of reasoning mechanisms and data structures that figure in communication among intelligent agents, and that work together to make implicature possible. These features should be independently motivated by linguistic and philosophical considerations, and should be theoretically central. I wish to nominate the following implicature-enablers: *speaker meaning as a sort of coordination-oriented intention*, *accommodation*, and the notion of *the conversational record*. None of these ideas is new; they have all been introduced by philosophers in the course of foundational projects having to do with pragmatics, and have been found useful by other philosophers, and by linguists.

#### 4 Accommodation as a Stumbling Block

Pointing out the importance of language use, as I have done, is fairly easy. Building a theory of language use is less easy. It is not enough to have a

large body of data, and methods for acquiring the data; discourse analysis, for instance, provides this. (See, for instance, Brown and Yule 1983.) And it is not enough to spin out theories. The problem is how to put theories in a productive relation with evidence.

If we take syntax as a model of what we should look for in a theory of pragmatics, we should seek to relate theories to linguistic evidence by way of a robust body of generalizations and tests. Many syntactic generalizations (such as "A tensed verb of an English clause agrees in number with the subject of the sentence") are intuitively appealing, are well supported by grammaticality judgments, and—though they are to a large extent independent of the details of particular theories—can be explained by syntactic theories. These properties enable informally stated generalizations to knit theories to evidence.

In pragmatics, the plausibility of informal rules is diluted by the fact that they are routinely flouted. Because of this, well-motivated generalizations not only will have exceptions, but many of these exceptions will be so flagrant as to seem to undermine their ability to serve as linguistic generalizations. To illustrate the phenomenon, I'll give three examples.

1. *Presuppositions of counterfactuals*. Considering the merits of Stalnaker's claim that when *A* semantically presupposes *B* it will be inappropriate to assert *A* unless *B* is presumed true, Lauri Karttunen says the following.

For example, consider counterfactual conditionals. It seems clear that a sentence like "If Bill had a dime he would buy you a coke" can be a felicitous utterance even in contexts where the truth of "Bill does not have a dime" is not taken for granted by anybody else but the speaker himself. One can utter a counterfactual conditional "in good faith" with the intent of informing the listener, among other things, that the antecedent clause is false. (Karttunen 1973, 170–171)

Suppose that Karttunen is right about the semantic presuppositions in this case. Even so, does the example provide an argument against Stalnaker? Suppose that in fact there is a primary pragmatic rule that 'If Bill had a dime he would buy you a coke' should not be asserted unless it's presumed that Bill doesn't have a dime. Then the sentence should also be usable to assert that Bill doesn't have a dime; the primary rule can be flouted to achieve this effect.

2. *Factive presuppositions*. To take a similar example, imagine that you see the following sign posted in a hotel elevator.

We regret that, due to renovations, our swimming pool will be closed to guests during the week of February 3.

The purpose of this sign is to inform guests that the pool will be closed. This would be a violation of Stalnaker's rule, on the assumption that 'regret' is semantically factive. That is, the rule that what is semantically presupposed should already be presumed true is flouted by the sign.

3. *Deictics*. Finally, here is a plausible rule about the use of deictic expressions: Do not use a deictic expression in a context in which the reference is not made clear.<sup>9</sup> As a corollary, you would not expect a conversation to begin with a sentence like 'Three of the others had called her'. But this rule is often flouted in the opening sentences of literary works. The first sentence of Tom Wolfe's *The right stuff* is a particularly striking specimen of the phenomenon. (And, partly because it is a violation of the rule, it is an example of good, vivid prose style.)

Within five minutes or ten minutes, no more than that, three of the others had called her on the telephone to ask her if she had heard that something had happened out there.

The mechanism that figures in this sort of rule violation is *accommodation*, which is described in detail in Lewis 1979. Most generally, accommodation consists in acting to remove obstacles to the achievement of desires or goals that we attribute to others.

I am accommodating you, for instance, if I open the door when I see you approach it with your hands full of packages. We can gain important insights into pragmatics by noticing that accommodations can affect conventions, norms, or data structures, as well as states of nature. A hotel in the Caribbean might accommodate a German tour by accepting German currency, and the U.S. Internal Revenue Service accommodates the public by allowing income figures to be rounded off to the nearest whole number. These examples of accommodation involve temporary or permanent modifications of rules, but accommodation can be spontaneous and ad hoc; discovering that I am a dollar short, a shopkeeper may decide to mark off the price of the item I have bought.

The case in which a shopkeeper regularly marks off his goods for various ad hoc reasons is different from the case in which the goods have no price at all, even though the cash register receipts may be the same for the two cases. In the one case there is a rule established by a marked price, in the other there is not.

There is a natural temptation to say that a linguistic rule can't be a rule if it is regularly flouted. I have come to believe that progress in pragmatics depends on resisting this temptation, and allowing for rules that remain significant even when they are flouted. In his work on implicature, Grice appeals implicitly to such hidden significance; he postulates rules of usage (called maxims), but explains many phenomena by supposing an effect to be achieved by flouting these very rules. Though flouting creates method-

ological difficulties, I would like to see how far we can get in pragmatics by taking the mechanism behind flouting to be fundamental.

### 5 *Accommodation as a Starting Block*

I have been leading up to the point that to accommodate accommodation in pragmatics, we must relax the conception that prevails in syntax of rules or generalizations. If there is such a thing as accommodation, genuine pragmatic rules will be routinely violated, and these violations will not seem anomalous; they will achieve an effect that seems natural and normal. Unfortunately, this relaxation will loosen the connection between theory and evidence. Since rules can regularly be flouted, we can legitimately ignore certain counterexamples to rules; in fact, we can even take instances in which a rule is flouted to provide support for the rule.

In his methodology of science, Aristotle stresses that the accuracy of an inquiry depends on what the subject matter will bear. We should not expect biology to be as accurate as astronomy, or politics to be as accurate as biology. And the best indicator of the accuracy of a science is the extent to which its rules hold for the most part rather than universally. One conclusion that we could draw about pragmatics is that it is less accurate than other cognitive sciences, and in particular that its relation to linguistic evidence is weaker than what prevails in syntax.

Worse still, in pragmatics we can take instances in which a rule is flouted to provide evidence of a sort for the rule. We can argue like this: "There must, in our society, be a rule against hanging up a telephone without closing the conversation, or people would not violate this rule to achieve an effect of snubbing."

Consider the adage "The exception proves the rule." Pedants love to tell you that in this saying 'prove' has its root meaning 'to test', though ignorant people often misconstrue it to have the meaning 'to justify'.<sup>10</sup> No doubt the pedantic point is etymologically correct, and if you are thinking of rules like Ohm's law, it is hard to see how anyone could be wrong-headed enough to subscribe to this proverb. But in the case of social rules, I think we can see some point to the popular misinterpretation of this "linguistic fossil."

I am not recommending that we should declare pragmatics a methodological disaster area in which we encounter rules that have many exceptions, and which in fact are supported by some of these exceptions. I do say that accommodation is a fundamental mechanism in pragmatics, one we simply can't ignore. Unfortunately, it creates methodological problems that we need to work to overcome.

Our best hopes for fostering good interactions between evidence and theory in pragmatics, it seems to me, lie in concentrating on underlying

reasoning mechanisms, and on adopting an interdisciplinary approach that spreads the sources of evidence wider than is common in linguistics. If such an approach is successful, perhaps generalizations will nevertheless have some contact with evidence, though this contact may be (even) less direct than what we find in syntax. We might claim, for instance, that the effect of accommodating a rule is always to swap the initial context for one that would have made the utterance appropriate. There is something to this; but the claim remains weak and attenuated unless we can turn this vague "would have made appropriate" into some more well defined operator on contexts. In general, there will be many different ways to adjust the context to make an utterance appropriate; how do we choose the right one?

We also have to look for ways of using indirect evidence, of spreading the evidential support for a pragmatic theory as broadly as possible, and in particular of finding unreflective linguistic processes that relate to pragmatic theory. I think there is little hope that we will find direct tests for distinguishing cases where accommodation is involved from cases in which it isn't: a special tone of voice, for instance, that comes into play when accommodation is invoked. But, for instance, work on intonation such as that described in Pierrehumbert and Hirschberg (chapter 14 of this volume) could be helpful in developing a theory of context. We have to hope that work of this sort will converge to provide the relation to evidence that a successful pragmatics will require.

Another way of constraining theories is by insisting that they be related to experimentation with computer programs. Examining the performance of a program is a powerful way of relating theories to linguistic evidence, and certainly is very different from either the introspective methods that are common in syntax or the more sociological methods that are used in discourse analysis. Moreover, this demand forces theories to be explicit, which is an extremely valuable constraint in pragmatics; and it imposes efficiency requirements as well. Work of the sort reported in this volume shows that theories of language use *can* be implementable. To keep ourselves honest we should now insist, I think, that a pragmatic theory *must* be implementable to be taken seriously.

Concentrating on accommodation means shifting to reconstructed reasoning that underlies utterances. And it suggests that certain reasoning processes, such as intention recognition and cooperation, are central. Successful accommodation requires that we first recognize someone's intention to achieve a goal, and then establish goals of our own that will assist in achieving this goal. These reasoning mechanisms have become prominent in computational thinking about cooperative dialog as well. To a certain extent, this thinking is quite independent from any philosophical or linguistic influence; obviously, for instance, an automatic help system (for, say, the UNIX operating system) will have to guess at the user's purpose when an

incorrect command is formulated, and may then even go on to implement the conjectured command. Other computational work, such as that of James Allen, Philip Cohen, and Ray Perrault, not only converges on the reasoning mechanisms that are involved in accommodation, but is influenced by some of the same philosophical work that influenced me. See the bibliography below for references to this work.

## 6 The Conversational Record

At this point I will begin a more systematic discussion of the components of the pragmatic theory that I am trying to articulate, beginning with the conversational record.

When philosophers discuss rules they think of games, and David Lewis' game is baseball. Although it's hard to find examples of rules of accommodation in baseball,<sup>11</sup> this choice of games helps to make vivid the idea of *scoreboard kinematics*. The baseball scoreboard is a particularly garish way of keeping track of public information about the game at hand.

Similarly, let's imagine a public, evolving representation of the state of a conversation: the *conversational record*. Such representations have been invoked by pragmatic theorists under many names, and for a variety of purposes. The idea can be found in Stalnaker 1972, an early and fundamental contribution to pragmatic theory; and it is developed in Stalnaker's later works (see the bibliography below for references). It also appears in logically inspired theories of temporal reference and anaphora, such as Hans Kamp's discourse representation theory and Irene Heim's file change semantics,<sup>12</sup> and in linguistic work on deixis.<sup>13</sup>

Like the baseball scoreboard, the conversational record will be a bundle consisting of many components, some of which are more stable than others. The baseball scoreboard must record whether there is a runner on first base; it needn't record the batting average of the first baseman. What sort of information is stored on the conversational record?

For one thing, the conversational record should represent information *that is public*, that can be supposed to be available to all the conversational participants. This means that private memories, intentions, and beliefs of participants can't be assumed to be part of the conversational record unless steps have been taken to explicitly enter them. And these steps must conform to procedures that facilitate the publicity of what is entered on the record. I will have more to say about this constraint of publicity in the next section.

Also, just as linguistic evidence is needed to justify a syntactic structure, postulated elements of the conversational record should be supported by evidence from discourse. In general, the form of the argument is that the discourse couldn't develop coherently unless the relevant information were

entered on the conversational record. In Stalnaker 1975a, Stalnaker gives as a reason for adding presumptions to the conversational record the fact that utterances like "John has children and all of his children are bald" are acceptable, whereas ones like "All of John's children are bald and he has children" are not. In Heim 1982 and Kamp 1981, broad linguistic support can be found for putting a set of "mentioned items" on the record. Fillmore 1975 gives reasons for adding a "home base" to the record, to account for words like 'go' and 'come'. In Lewis 1979 there is a long list of elements that might well be included in the conversational record, but without any linguistic argumentation for including these elements.

More needs to be done to consolidate this work and to find plausible ways of reasoning about the conversational record—bearing in mind that the evidential situation is complicated by the presence of accommodation and that the conversational record itself is a fairly rarefied theoretical construct, that is not closely connected to linguistic evidence by simple procedures.

The idea of an evolving data structure that influences and is influenced by reasoning processes is familiar from computer science. The intentional structures and focus structures proposed by Grosz and Sidner (in Grosz and Sidner 1986 and in chapter 20 of this volume) are perhaps good examples of structures proposed for computational reasons that would fit naturally on the conversational record. At least, the discourse phenomena that these structures are supposed to underlie—anaphora and topic management—are things that should be public if a conversation is to go well.

I used the word 'perhaps' because Grosz and Sidner's work doesn't make it entirely clear how their structures can be mutually recognized by the participants in a conversation. But the key to clarifying this matter is to discover cues and constraints that can serve to effectively communicate these changing structures to the participants in a discourse. Progress toward implementing their proposals in a relatively general setting would depend on providing such constraints.

The information manipulated by plan-based approaches such as that of Litman and Allen in chapter 17 of this volume and Allen's earlier work is only partially assimilable to the conversational record; the reason for this is that these approaches depend on recognizing intentions and beliefs, and many intentions and beliefs that are recognized by some of the participants in a conversation aren't mutually recognized, and so are not public from a conversational standpoint. Consider the following two cases: (1) Standing by your parked car, you say "I'm out of gas" to me as I pass by; (2) Standing by your parked car, you ask me "Where can I find a phone booth?" We can imagine that in both cases the stage setting is the same: as I approached I heard your engine turning over without starting, and noticed an out-of-state plate. You emerged from the car as I came up. In

both cases I can infer that you intend to find a service station; but in the first case the intention is public, in the second it is not.<sup>14</sup>

If (as in the applications that advocates of the plan-based approach have in mind) we will be interested in producing helpful responses, we will in general be interested in using whatever conclusions we can reach about our interlocutors' beliefs and intentions, regardless of whether these are public or not. Intentions and beliefs that are *not* public will be especially important in applications of the sort discussed by Pollack in chapter 5 of this volume, where the task is to identify mistaken beliefs.

Any belief can affect our discourse. (For instance, it may be the topic of conversation.) So a theory of discourse that is continuous with linguistic methodology must limit the data and reasoning processes that are internal to the subject. I am suggesting public data and mutually recognizable processes as a plausible circumscription principle. The main test of this proposal will be its scientific usefulness; but it also has a kind of general plausibility. Mutually believed data, and procedures that contribute to what is mutually believed, are plausible candidates for obtaining a linguistic focus because there are very general considerations connecting mutual belief with convention, and convention with language; see, for instance, Lewis 1969.

There are few intelligent tasks that are merely linguistic, so it isn't surprising that serious applications in artificial intelligence involve both linguistic and extralinguistic reasoning. However, it serves some computational purposes (transportability, in particular), to separate these functions in developing intelligent systems. It remains to be seen whether a separation between domain planning and plan recognition on the one hand, and discourse planning and plan recognition on the other, is appropriate and helpful in artificial intelligence.<sup>15</sup> If it is, the theory that I am advocating here could be seen as the "linguistic" component of a discourse system. Current discourse systems represent the beliefs of the speaker and hearer, but do not attempt also to represent mutual belief. To achieve the separation of components I have in mind, the speaker's and hearer's beliefs would have to be distinguished from the ad hoc collection of mutual presumptions that figures in the conversational record. And some "beliefs" in the earlier work would appear as mutual presumptions. In some cases at least, this adjustment would be in the spirit of the work; what is represented as belief in much of the work is better construed as mutual supposition. Perrault's work in chapter 9 of this volume is a very good example of this.

For my purposes in this paper, the most important component of the conversational record is a structure *P* that determines the *presumptions*: the things that are supposed, or established, at a given stage of the conversation. I'll say that *P* ⊢ *A* when *P* yields the conclusion *A*. These presumptions are perhaps best thought of as a kind of shared memory or



common database that the participants construct for the purposes of the conversation.

Of course, presumptions will in general be modified in the course of a conversation. I'll suppose that for each proposition  $p$  there is an update operator  $\mathcal{A}_p$  on the presumptions that gives the result of updating the presumption structure by adding the proposition  $p$ .  $\mathcal{A}_p(\mathbf{P}) \vdash p$ , and in many cases  $\mathcal{A}_p(\mathbf{P}) \vdash q$  if  $\mathbf{P} \vdash q$ . However, update is a nonmonotonic operation, so that we can have  $\mathbf{P} \vdash q$  but  $\mathcal{A}_p(\mathbf{P}) \nvdash q$ ; also we can have  $\mathcal{A}_p(\mathbf{P}) \vdash r$  but  $\mathcal{A}_{p \wedge q}(\mathbf{P}) \nvdash r$ , where  $\wedge$  is propositional conjunction. Since the conversational record is public, the update operator must be public also. I assume that the reasoning mechanisms that operate here are essentially the same as those that operate in the suppositional or conditional reasoning of a single agent. (This reasoning is not particularly well understood, but fortunately what I will have to say here will not require a detailed theory of it.)

There is a close connection between *assertion* and the update operator; in fact, it is very like the connection between buying something and obtaining it, only perhaps more intimate. The goal of asserting a proposition is updating the conversational record with the proposition; and asserting a proposition is the standard way of updating the conversational record with the proposition.

Assertion as a speech act can't be *identified* with the update operator, because a speaker can effect an update without asserting anything. Though he made an update occur, Herod didn't assert anything when he caused John the Baptist's head to be brought in on a platter. But, just as obtaining is the point of buying, update is the point of assertion.

It is really a methodological issue whether we should expect the conversational record to contain enough information to characterize the goals of all or most speech acts. The goals of insulting and apologizing, for instance, may force additions to the conversational record that could not be supported by linguistic argumentation. But even if our criteria for additions to the conversational record are fairly strict, I think we would have the materials for characterizing the goals of many general types of speech acts.

At this point, assertion looks fairly simple: the general picture is just this.

Background		Input		Shifted
presumptions	+	proposition	$\Rightarrow$	presumptions
$\mathbf{P}_1$		$p$		$\mathbf{P}_2$

But this simplicity is misleading; it disappears as soon as we talk about the *act* of assertion, rather than just its goal. If we consider how conversants go about achieving updates of the presumptions, things become more complex because of the strategic planning of speech acts. A speaker in a context  $c$  in which  $\mathbf{P}_1$  is presumed may want to achieve a context in which  $\mathbf{P}_2$  is

presumed, and so will have the goal of getting  $p$  asserted. One way to accomplish this might be to utter a declarative-mood sentence expressing  $p$ . But there may be other ways: instead of saying to my wife "I'll need the car this afternoon," I say "I didn't tell you that I'll need the car this afternoon." Strategic speaking of this sort, of course, is one reason why we need a theory of implicature.

### 7 Public Constructs

The idea of explaining the goals of acts of meaning in terms of the beliefs and intentions of conversants goes back at least to Grice 1957, in which the point of a meaningful assertion is the creation of a belief, and the point of a meaningful imperative the creation of an intention in the hearer. This idea has been reproduced in the computational literature, and in fact can be found in most of the work in which planning is applied to discourse; Allen 1983 is a good example.

According to this picture of things, the participants in a conversation are attempting to work on each other. I have said that I would like to replace this with a slightly different picture, according to which they are working together to build a shared data structure. What I have in mind is like a group of people working on a common project that is in plain view. For instance, the group might be together in a kitchen, making a salad. From time to time, members of the group add something to the salad. But it is assumed at all times that everyone is aware of the current state of the salad, simply because it's there for everyone to see.

Thinking of things in this way is useful philosophically; for instance, see the discussion of speaker meaning in section 9, below. I hope, of course, that it could be useful computationally as well, but at this point the hope is untested.

I grant that in most cases that we'd want to describe as "conversation," there won't be anything like a literally shared memory. Normally, there will be no shared conversational record; each conversant will have its own representation of the record. But it is normal for conversants to assume that the representation is shared, for each conversant to treat its representation of the record *as if* it were a public object. Unless danger signals are perceived, a conversant won't distinguish between its representation of the record and its interlocutors' representation. If we were designing programs for conversing systems along these lines, the aim would be to have things work out so that the representations will start out coordinated, and coordination will be preserved by transformations, so that at all stages the representations of the record will match.<sup>16</sup>

In its simplest form, this would merely be the problem of getting two communicating database programs to exchange information in such a way



that if they start out with the same information they will end up with the same information—not a difficult task, if the communication channel is large enough, and is error-free.<sup>17</sup> But the task becomes more complicated if the systems can plan strategic speech acts. To maintain coordination of the record, the two systems must plan these acts so that they will be successfully decoded; this means that they must plan communication acts in such a way that their plans will be recognized. (A good plan to put something on the conversational record must be a plan for entering the information publicly, in view of the coordination constraint.) This way of looking at the problem could conceivably eliminate some of the logical complexity that finds its way into belief-based computational models of discourse.

To a large extent, the process I indicated would depend on the programs not only employing default reasoning in the recognition of intentions, but on their being able to reflect on this reasoning in planning utterances; that is, these programs must plan their utterances so that their interlocutors will default to the right interpretation. Though, as I say, I can't yet make any claims about the computational advantages of the conversational record, it in fact seems from Perrault's recent work that default reasoning does simplify things considerably. (See chapter 9 of this volume.)

Using presumption rather than belief would help the motivation of Perrault's rules: for instance, by eliminating the need to postulate the principle that hearers believe what speakers say. This is not a plausible principle of either conversation or epistemology. However, it is a reasonable principle of conversation that unless someone objects, what a speaker asserts is added to the conversational record.

In saying that assertion does not aim intrinsically at belief, I have driven a wedge between communication and belief. Of course, this also divorces communication from action—and perhaps one motivation of the belief-based approaches was to connect the two. On the view I'm advocating, what is the connection between assertion and belief? Well, the conversational record is like a courtroom record; it is evidence, but we may not choose to believe all of it. What we choose to believe from what we are told is a special case (but a maximally complex special case) of what to make of evidence. Moreover, the very notion of belief is an oversimplification of the reasoning phenomena. Rather than a single system of beliefs that we maintain and apply in all situations, we are capable of harboring a variety of systems of suppositions, which are activated and which we are more or less willing to act on according to contextual factors. (See Thomason 1987.) Conversational presumption is a case of ad hoc supposition, very like supposition for the sake of argument. However, supposition is usually serious rather than idle, and more often than not is credulous. In such contexts there is no practical difference between presumption and belief.

Consider the case of a congressman interviewing a witness at a hearing. What the congressman has to assume in order to keep track of the conversation and what he believes are two different things. Contrast this with a conversational system that is designed to be a total slave to its interlocutor; here, we can ignore the distinction between presumption and belief, and in other cases we might well ignore the distinction in order to simplify things.<sup>18</sup>

I don't have a very good picture of how the conversational record is initialized in human conversations. Topical or salient mutual beliefs normally go on the record; and so, the better acquainted the conversants, the larger the initial record. If the conversants' models of one another are badly out of phase, we may not be able to reconstruct a record.

The record can be changed in a variety of ways. At the most general level it can be changed either in response to conversational moves of the interlocutors, or by manifest, public occurrences in face-to-face conversations, which add to mutual beliefs. (Remember John the Baptist's head.)

The idea of the conversational record and its kinematics is one place where the research strategies of philosophy, linguistics, and computer science should blend fruitfully and reinforce one another. Philosophers can demonstrate the conceptual need for the idea and reflect on its foundations; linguists can show what things need to be built into the conversational record in order to account for linguistic phenomena; and using devices such as plan recognition, default reasoning, frames, scripts, etc., computer scientists can consider how to represent and augment the conversational record by means of feasible reasoning processes.

### 8 Plan Recognition

Through a process that is quite different from the methodology of either philosophy or linguistics (since it involves experimentation with the performance of computer programs) computer scientists working on discourse have come to see plan recognition as centrally important. Details on this subject can be found in many of the computational papers in this volume, and in the works cited in these papers.

But philosophical reflections on discourse by themselves can show that plan recognition is central from a general perspective. (Perhaps this point hasn't struck philosophers because plan recognition isn't familiar to philosophers as a category of reasoning.) I am leading up to the point that plan recognition is as important for understanding implicature as deduction is for understanding validity. For now, I want to illustrate the point for two notions that are connected to implicature: *speaker meaning* and *accommodation*.

### 8.1 Plan Recognition and Speaker Meaning

On a simple version of the account of speaker meaning in Grice 1957, a speaker *S* means *p* if *S* intends the hearer (1) to believe *p* and (2) to do so partly in virtue of the recognition of intention (1). For some reason, not much has been said in the philosophical literature about the concept that is invoked in (2)—“to come to believe *p* partly in virtue of recognizing intention *i*”—even though it is an indispensable part of Grice’s analysis, is conceptually complex and philosophically problematic in its own right, and does not appear to be definable in terms of concepts like belief and intention.<sup>19</sup>

It is fair to say that this formula of Grice’s is pointing squarely at the mechanism of plan recognition. Plans are wholes constituted of intentions, and we recognize intentions by fitting them into recognized plans.<sup>20</sup> To come to believe *p* partly in virtue of recognizing intention *i* is to come to believe *p* by recognizing a plan that involves *i*.

This fits in well with the general philosophical strategy of Grice 1957; one of the main goals of the paper is to account for speaker meaning in terms of concepts and capabilities that are not essentially linguistic. And there is nothing inherently linguistic about plan recognition; we are using plan recognition when we infer that someone is going to the airport on seeing them standing with luggage at the limousine stop.

The fundamental notions in Grice’s 1957 definition, then, are belief, intention, and plan recognition (though he doesn’t make the last of these explicit). Later, I’ll suggest that the definition can be improved by substituting the conversational record, planning, and plan recognition for Grice’s notions. Any good account of speaker meaning, however, should connect acts of meaning to planning and plan recognition, providing ways in which what we know about these reasoning mechanisms can clarify and inform the theory of speaker meaning.

### 8.2 Accommodation

Acting as if we don’t have a flat tire won’t repair the flat; acting as if we know the way to our destination won’t get us there. Unless we believe in magic, the inanimate world is not accommodating. But *people* can be accommodating, and in fact there are many social situations in which the best way to get what we want is to act as if we already had it. Leadership in an informal group is a good case. Here is an all-too-typical situation: you are at an academic convention, and the time comes for dinner. You find yourself a member of a group of eight people who, like you, have no special plans. No one wants to eat in the hotel, so the group moves out the door and into the street. At this point a group decision has to be made. There is a moment of indecision and then someone takes charge, asks for suggestions about restaurants, decides on one, and asks someone to get

two cabs while she calls to make reservations. When no one objected to this arrangement, she became the group leader, and obtained a certain authority. She did this by acting as if she had the authority; and the presence of a rule saying that those without authority should not assume it is shown by the fact that assuming authority involved a certain risk. Someone could have objected, saying “Who do you think you are, deciding where to go for us?” And the objection would have had a certain force.

Another familiar case, involving even more painful risks, is establishing intimacy, as in beginning to use a familiar pronoun to someone in a language like French or German. Here the problem is that there is a rule that forbids us to act intimate unless we are on intimate terms; and yet there are situations in which we want to become intimate, and in which it is vital to do it *spontaneously*, rather than by explicit agreement. If I find myself in such a situation, my only way out is to accept the risk, overcome my shyness, and simply act as if you and I are intimate, in the hope that you will act in the same way. If my hopes are fulfilled, we thereby *will have become* intimate, and it will be as if no social rule has been violated. A process of accommodation will have come to the rescue.

David Lewis elaborates many examples of accommodation—presupposition, permission, definite descriptions, coming and going, vagueness, relative modality, performatives, and planning—most of them having to do with linguistic rules. And he gives the following general scheme for rules of accommodation.

If at time *t* something is said that requires component *s<sub>n</sub>* of conversational score to have a value in the range *r* if what is said is to be true, or otherwise acceptable; and if *s<sub>n</sub>* does not have a value in the range *r* just before *t*; and if such-and-such further conditions hold; then at *t* the score-component takes some value in the range *r*. (Lewis 1979, 347)

No doubt you could implement rules of accommodation in this form.<sup>21</sup> But I think we can understand the process of accommodation better, and assimilate it more readily to other sorts of reasoning, if we think of it in terms of a combination of *plan recognition* and *cooperative goal adoption*.

In fact, I believe that accommodation is a special case of obstacle elimination, a form of reasoning discussed in Allen 1983. Obstacle elimination consists in (1) recognizing the plan of your interlocutor; (2) detecting obstacles to the plan in the form of certain false preconditions of subgoals belonging to the plan; (3) adopting the goal of making these preconditions true; (4) forming a plan to carry this out; and (5) acting on this plan. Step (1) is plan recognition. Step (3) is cooperative goal adoption. The other steps employ forms of reasoning that figure in noncooperative planning by isolated agents.

Opening a door for someone is a form of obstacle elimination. So is adding  $p$  to the presumptions when someone says something that presupposes  $p$ . The difference between the two has mainly to do with the social nature of the conversational record. In the case of the door, we simply don't have the practical option of acting as if the door were already open. In the case of the conversational record, to act as if the previous state of the record already involved the presumption  $p$  is to reset the record. The fact that changes in the conversational record can be made so effortlessly accounts in large part for the extensive role that is played by accommodation in conversation—at least in informal and noncompetitive conversation.

The principle behind accommodation, then, is this:

Adjust the conversational record to eliminate obstacles to the detected plans of your interlocutor.

If the term hadn't already been claimed, this could well have been called the *cooperative principle*.

### 9 Speaker Meaning

Ordinarily, we mean something by what we say. We can mean things without saying anything, and when we say something the meaning of the words we say need not be the same as what we mean; the literal meaning of our utterances can be displaced. We need an independent account of what it is for a speaker to mean something. Over twenty-five years ago, before he had developed the theory of implicature, Grice published such an account (see Grice 1957).

The paper does not explicitly say where the idea of concentrating on speaker meaning came from; my guess would be that it originated (at least in part) in a critical chain of reflection initiated by J. L. Austin's work on illocutionary acts, and some other philosophers' ill-conceived attempts to characterize meaning in terms of use. The paper presents a definition of what it is for A to mean  $p$  by doing  $x$ . Grice's account of what he calls "non-natural meaning" is thoroughly familiar to philosophers, but may not be as well known to linguists and computer scientists as his work on implicature. Here is a formulation of the definition that will serve the purposes of this paper:<sup>22</sup>

A means  $p$  by  $x$  in case A intends an audience to believe  $p$  by (in part) recognizing this intention of A's.

For instance, by looking at his watch A means (to B) that it's time to go in case A intends B to believe that it's time to go, and to come to have this

belief partly because B recognizes A's action as a sign of A's intention to communicate the belief that it is time to go.

Before suggesting a revision of Grice's definition, I want to return once more to the ingredients of the revision, and to suggest that communicating creatures with the capacity to accommodate each other and the capacity to reflect on this capacity will be able to deal in *strategic meanings*: in speaker meanings that differ from the literal interpretation of the message. That is, creatures with these abilities will implicate.

Part of the capacity to accommodate is the ability to recognize plans, at least in familiar domains in which one is able to plan effectively for oneself. By assumption, communication is such a domain for these creatures. Therefore, these creatures will be able to recognize plans to make  $p$  asserted, plans to update the record with  $p$ . Since they can accommodate, they can update with  $p$  as a result of recognizing such a plan. Since they are able to reflect on their abilities to accommodate, they must in some cases be able to form plans to make things asserted that exploit this communicative mechanism. But an utterance may reveal a plan to make  $p$  asserted without using language that literally expresses  $p$ .

Most conversational implicatures, I think, are meant;<sup>23</sup> and in fact the response "I didn't mean that" can always be used to renounce an implicature. Take one of Grice's examples: A says "Where does Smith live?" and B replies "Somewhere in the south of France," implicating that B is not able to specify the town where Smith lives. Here, it seems right to say that B *meant* that he didn't know what town Smith lives in.

One effect of the revision I'll propose of Grice's 1957 definition is that strategic assertion will be an instance of meaning. Like Grice, I offer this suggestion in an experimental spirit, though I have in mind a somewhat different arena of experimentation. In this spirit, the proposal is tentative. I hope that at least it is a good beginning.

To mean  $p$  is to intentionally reveal an intention to make  $p$  asserted through the hearer's recognition of the status of an intention or plan of the speaker's.

Notice that this is a considerable departure from Grice's 1957 definition; and it makes plan recognition central to meaning. In fact, it is very close to saying that to mean something is to communicate it through plan recognition. It is also much less convoluted than Grice's definition;<sup>24</sup> nevertheless, it gives much the same results as Grice's, over a healthy spread of cases.

It should be clear now that to make  $p$  asserted is to add  $p$  to the presumptions of the conversational record; perhaps by updating the presumptions only with  $p$ , perhaps by updating them with other propositions as well. And when I discuss examples, it will become clear that the intention may have either an "activated" status, or a "frustrated" status.

This is an account of "indefinite meaning"—of what it is for  $p$  to be a thing that is meant, to be one among many things that are meant. Just as a plan may have multiple goals, we often mean several things by a given utterance. For instance, in Example 3, below, A means by "There's a gas station around the corner" that there is a gas station around the corner and also that the station is open. Just as one goal may dominate a plan with multiple goals, we often identify one meaning as preeminent among the things that are meant on a given occasion; this is *what is meant* on that occasion, or *the meaning*. This "definite" sort of meaning is not easy to characterize, and neither Grice nor I try to define it.

I'll illustrate the definition with some examples, chosen for variety.

*Example 1.* A says "I'll need the car," meaning that A will need the car. In this case the plan is to add a proposition  $p$  to the record by uttering a sentence that expresses  $p$ . The intention that is meant to be recognized is the very intention to assert  $p$ . I am in effect agreeing with Grice that when  $p$  is meant, the assertion may be intended to "take hold" partly because of the recognition of the intention to assert it. That is what happens in this case. I'm not sure, though, that this is so in all cases of meaning  $p$ .

Example 1 is a humdrum case, of course, but humdrum cases can be important, and actually these simple cases are problematic for Grice's 1957 definition, because they require the postulation of far-fetched hidden intentions underlying automatic, unreflective actions. Of course, postulating such intentions is a fundamental tactic not only of Grice's methodology, but of much work in cognitive psychology. If we subscribe to this methodology, I think we can say that the speaker intends his plan to make  $p$  asserted to be recognized. And in general, a hearer who did not recognize an intention to make  $p$  asserted would not update the record with  $p$ .<sup>25</sup> My own methodology also depends on ascribing far-fetched plans to conversational agents, plans that these agents might not reflectively acknowledge. So at least, I think I can make the following claim: if this case fits Grice's 1957 definition, then  $p$  would not become asserted without the recognition of the intention to assert it.

*Example 2.* A looks at his watch, meaning that it's getting late. Here A plans to make the proposition that it's getting late asserted. In some cases of this sort, at least, A intends the recognition of his intention to assert  $p$  to figure in the assertion mechanism. It is these cases that seem to be the ones that we would intuitively think to be instances of meaning that  $p$ . This is exactly what Grice noticed in motivating his 1957 definition. If, for instance, A's plan were rather to make the hearer aware that A thinks it's late and so to offer to leave, we would be less likely to call it a case of meaning anything. The account that I am urging agrees with Grice in claiming that in this second case nothing is meant either. But the reason for this is rather that A would have no practical reason to consider this to be a way of

asserting  $p$ . I am supposing that the intentions involved in meaning must (like any intentions) be practical; and so A can't mean  $p$  unless A has a more or less sensible plan for putting  $p$  on the public record. What we have said in this case about A's intentions does not suffice to do this, because it doesn't sustain the creation of a mutual presumption. For instance, A is not, so far as we have described this case, intending that the hearer presume that A presume that the hearer presume  $p$ .

*Example 3.* A says to B "There's a gas station around the corner," meaning that the gas station is open.<sup>26</sup> I am imagining a context in which B plainly needs gas, in which A's discourse plan is to recommend to B a plan for getting gas, and in which B can easily recognize this discourse plan.

Ordinarily in such a case, I'd want to say that A intends the proposition that the station is open to be added to the conversational record. And in general, when a speaker openly recommends a plan, he (equally openly) invites the presumption that it is feasible: that the preconditions of the plan that are not within the agent's control are true (or will be true at the appropriate time). This seems a reasonable principle of conversation, as plausible as the principle of deliberation that in planning from an intended goal to a plan involving subgoals, intention will transfer to the subgoals. So in recommending a plan to get gas at a station, one is openly inviting the presumption that the station is open.

In this case, however, the primary discourse intention is to add to the record the proposition that there is a gas station around the corner; the intention to add the proposition that it is open (if the intention exists) is secondary. (This is what makes the assertion an implicature.) And unfortunately, the notion of what counts as a secondary intention is rather loose. When I buy a state lottery ticket in Pennsylvania, do I intend to contribute to the fund for the elderly that I know it supports? It is hard to say.

But secondary meanings are also hard to pin down; and if my definition is right, this is because they involve secondary intentions. In the gas station case, perhaps the best we can hope for is indicative evidence, such as the fact that in this case it would be natural for B to reply "How long will it be open?" We can also see that it is A's intention to recommend a plan *to get gas* that serves here to create the meaning.<sup>27</sup> Notice, for instance, that if B were on foot and had said to A "I need to make a phone call," then this same verbal response in A's mouth would not mean that the gas station was open—at least in cases where public telephones are often outside of gas stations.

A noteworthy point. We could have stated the proposed definition of meaning like this:

To mean  $p$  is to intentionally reveal an intention  $i_1$  to make  $p$  asserted through the hearer's recognition of the status of an intention or plan  $i_2$  of the speaker's.

In the first two examples,  $i_1 = i_2$ . But in Example 3,  $i_2$  is the intention to propose a plan for getting gas. For Grice (if we equate my "presumption" and his "belief" for the moment), the intention to make  $p$  asserted/believed must be part of the reason for the assertion taking hold. I am not sure how to settle this point definitively, but in the gas station case it seems to me that the speaker does mean (in many normal cases) that the gas station is open, and that it is more plausible to say that the assertion is intended to take not because of recognition of the intention to assert that it is open, but rather simply through recognition of the speaker's plan to assist the hearer with his domain problem. Certainly there is a direct chain of reasoning from recognizing this dominant plan to the presumption that the station is open, and there is no need to postulate as figuring in this process the recognition of a separate intention to assert that the station is open.

This is a significant organizational difference between my proposal and Grice's 1957 definition. It may expose me to counterexamples, but so far I haven't discovered any.

*Example 4.* B says to A: "I need to send a letter to Smith; where is he?" A says "Somewhere in the south of France," meaning (among other things) that A does not know what city Smith lives in. Like the previous example, this is a case in which the speaker's purpose is to assist the hearer with a domain plan. However, the response clearly is inadequate for B's purposes. Again, I have no definitive way of showing that A intends to add to the record the proposition that A doesn't know where Smith lives. It is indicative that B could say "Well, do you know where Jones lives?" with heavy stress on 'Jones'. Also, in most cases A would want it understood that he is cooperating. I'll assume that A intends the proposition to be added.

In this case, I want to say that A is disclosing two plans: a discourse plan to answer B's question (on which A is acting) and a frustrated domain plan to help B mail the letter. Recognition of the frustrated intention to tell B the address, and identifying A's lack of knowledge as the obstacle to fulfilling the intention, is the reasoning process that is meant to enable the proposition that A does not know what city Smith lives in to be added to the presumptions.

*Example 5.* A says to Johnny, in the presence of B: "Now Johnny, tell B you dug up his bush." Here A does reveal an intention to make a proposition asserted, but also issues a command that will make the proposition asserted without reference to that intention. In cases such as this, A may or may not intend  $p$  to be added to the record (in part) through the recognition of an intention or plan of A's. If A does so intend, then according to the definition A means  $p$ ; if not, A is trying to get Johnny to mean  $p$ , without meaning  $p$  himself. This result seems right to me. In cases where the alternative presumption-recording process is particularly vivid, like the head-on-the-platter case, it is more difficult to imagine the speaker meaning anything.

As in Example 2, practicality comes into play here. The reason why the intention is implausible is that an intention to have recognition of the intention to assert  $p$  figure in making  $p$  asserted would simply not be very realistic in such cases. Except in lapses, like the case in which the Reverend Spooner said "Leave by the town drain," meaning "Leave by the down train," I assume that our discourse intentions are normally practical.

*Example 6.* A and B are planning to go shopping. They need \$50. B says "I have \$20." A says "I have \$40." A does not mean by this that A has more money than B.

There is no point in making things easy for ourselves by restricting the application of a theory of speaker meaning to positive instances. Negative cases are particularly good ways to discover weaknesses in theories.

This example illustrates a general point about intention-based accounts of speaker meaning; for such accounts to work, the intensionality of meaning has to be supported by an equal intensionality in intention. If there is no difference between intending to make  $p$  asserted and intending to make  $q$  asserted, then there can be no difference between meaning  $p$  and meaning  $q$ .

Unfortunately, there is no direct way to test for subtle differences of intentions of this sort. On the one hand, it does seem in this example that A meant he had \$40, and didn't mean—even among other things—that he had more money. On the other hand, we'd expect A to say "Yes" if we performed the experiment of asking him "Did you mean to say that you had more money than B?" All this shows, I think, is that 'Did you mean to say' is used to query people about the consequences of what they said, without distinguishing these clearly from what they said. More generally, 'Did you mean to do' does not distinguish between intentions and foreseeable side effects, which is exactly the distinction we need to make here. To achieve some progress in these negative cases, it seems that we will need to say more about the intensionality of intentions in general, and of intentions to assert in particular, and also to find tests that extract more reliable conclusions from the evidence. The philosophical background, of course, suggests that it won't be easy to make progress on these matters.

Though the example points to weaknesses in the account, it is at least not a counterexample. The definition tells us that A didn't mean that he had more money than B if he didn't intentionally reveal an intention to assert that he had more money than B through B's recognition of an intention or plan of A's. And this seems plausible to me. If we change it to a case in which A does intentionally reveal such an intention (say, A and B have been arguing about who has more money), we change whether A means that he has more money.

On very crude grounds—simply asking oneself whether it would be natural to acknowledge an intention to assert something—I think that the

judgments I have made here about whether *p* is intended to be asserted are plausible. But we badly need to be able to say much more about the reasoning mechanism, to show how intentions to assert are recognized. (It follows from what I said before that there must be coordination between the cases in which practical intentions of this sort exist and cases in which their recognition will occur.) In effect, what is needed is an account of discourse plans and how they are recognized. I hope that work such as that of Litman and Allen will help to fill in these details (see chapter 17 of this volume).

The major strategic difference between this revision and Grice's approach is the substitution of the presumption dimension of the conversational record for belief. As I indicated, this eliminates a class of glaring counterexamples to Grice's 1957 definition: answers to examination questions, where the hearer already believes the answer, or statements by a witness who has no hope that what he is putting on the courtroom record will be believed by the audience. Also, I am hoping that the fact that the conversational record must be public will absorb much of the complexity of Grice's meaning account.

Openly meaning to assert *p* is very close to meaning *p*. This is an (intended) consequence of the revised account. It is part of the definition that one must mean to assert *p* when one means *p*. Conversely, if *A* means *p*, *A* is intending to have the assertion take hold through recognition of the status of an intention of *A*'s. And ordinarily when one openly intends to assert *p* (that is, when one intends to assert *p* in a way that aims at making itself recognized), there is no independent mechanism that will equally well cause *p* to be asserted. Usually we only intend to do things that need to be done. The main class of counterexamples are cases in which there is a separate mechanism, but the mechanism is not independent: the cases that I can think of are like Example 5. These do happen, but not all that frequently. And besides, most of them are cases in which it is not unnatural to say that the proposition is meant: for instance, it is plausible that I mean that you ate my candy when I say "Come on, admit that you ate my candy."

#### 10 The Update Mechanism

It should be clear that I am aiming for a close relation between meaning and the mechanism for updating the conversational record. In the declarative case, the central mechanism is simply *the recognition of intentions to assert*—that is, the recognition of intentions to perform the update in question—and this is a special case of accommodation.

In fact, I am proposing as a general rule that whenever an intention to assert *p* is recognized, the record is updated with *p*. It is this rule, I think,

that forms the core of the insight that conversation is in some sense cooperative.

This principle, of course, is a default rule; the update will in general not take hold if some appropriate participant in the conversation objects. Also, the principle is subject to felicity principles similar to those noted in Austin 1962b. For instance, the assertion may not take if the speaker is not authorized to make it: in ritual circumstances, as when someone speaks out of turn in a courtroom, and also in informal cases when someone simply speaks out of turn. And it may not take if the proposition is grossly off the topic.

The picture of update that I painted earlier is complicated by the fact (which shows up clearly in some of the examples that I presented in section 9, above) that speakers often kill many conversational birds with one stone; they intend at once to assert a number of propositions. In such cases, we can assume that the update is achieved by adding the conjunction of the propositions that are meant (though of course there may be differences in salience among the asserted propositions). However, we can expect that the problem of plan recognition in discourse will be complicated by this multiplicity of utterance functionality, since the recognition of multiple plans appears to be a much more difficult problem in general than that of single plans.

As I have said, for the record to remain public, the update process must remain coordinated; there must be a convergence between the things that are meant and the things that are recognized as meant; this will depend on successful modeling of the hearer by the speaker.

#### 11 Conversational Implicature as Based on Accommodation

I have never been convinced that it is very useful to use the term 'conventional implicature' for the phenomena that Grice aggregates under this label.<sup>28</sup> So in what follows I'll simply talk about implicature, meaning something close to what Grice meant by conversational implicature. And I will confine myself, as before, to cases of assertion.

Implicature is chiefly to be classified with straightforward cases of assertion. Roughly, an implicature is something that in general we would agree is said and is meant, but that is meant indirectly, so that the speaker could (perhaps disingenuously) exercise an option to disavow it, saying "I didn't say that."<sup>29</sup>

Consider again the examples of section 9. In Example 1, something (the proposition that it's getting late) is meant, but not implicated. In Example 3, *A* says and means (but does not implicate) that there is a gas station around the corner; *A* also means and implicates that the station is open. In Example 4, *A* says and means (but does not implicate) that Smith lives in the south

of France; A also means and implicates that A does not know what city Smith lives in.

Largely because the notion of implicature is negative (roughly, it intersects what is meant with what is *not* said explicitly), I think its main theoretical importance is to call our attention to the fact that many things are meant without being explicitly said. Of more importance are the mechanisms that allow meaning to outstrip what is explicitly said: mechanisms like plan recognition, goal sharing, and accommodation. I have tried to write this paper to focus on these underlying mechanisms, and this must have made it seem that I was never going to get to the point. Now I would like to try to apply what I have said to implicature.

Part of what I am claiming is that accommodation is a common underlying *cause* of implicature, a cause that results in an assertion-like transformation of the presumptions.

But this effect can take various forms—in particular, it may involve either an *accommodated revision of the background presumptions*, or an *accommodated revision of the input proposition*. This distinction gives rise to two kinds of conversational implicatures: *background implicatures* and *assertional implicatures*.<sup>30</sup>

An example of a background implicature is the case in which I say “I regret that I can’t help you,” saying and meaning that I regret that I can’t help you, and meaning that I can’t help you. (This is a case in which the implicature is both entailed and presupposed by what is said.) Another rather different example, which I’ll discuss in more detail later, is “Arnold doesn’t know that his shoelace is untied,” used to implicate that Arnold’s shoelace is untied. In both of these cases, I think, something is said that would be inappropriate if the current presumptions didn’t involve the implicated proposition. The actual current presumptions may not involve this proposition, and the speaker makes the utterance with the intention that the hearer will accommodate it by amending the current presumptions, as well as adding the proposition that is directly asserted.

In a way, the term ‘background implicature’ is misleading; a background implicature, as the above examples show, may be the most important and newsworthy proposition that is conveyed by an utterance. In general, matters having to do with things like newsworthiness are orthogonal to matters having to do with mode of conversational presentation.

A case of an assertional implicature (one that I’ll soon discuss in more detail) is the example in which I say “I didn’t tell you I’ll need the car” to my wife, implicating that I’ll need the car. Here, the proposition that is literally asserted can’t be used for update; it is already presumed. Rather, it is the implicature that is asserted. And if Grice is right about irony, some uses of this turn of speech will be assertional implicatures: for instance,

“Television is doing wonderful things for the ability of the American college freshman to write.”

### 12 An Example

Let’s consider again the case in which I say to my wife “I didn’t tell you I’ll need the car this afternoon.” When I say this, of course, the main thing that I want to get across is that I’ll need the car; and I do manage to get this across. But what I literally say is that I didn’t tell her I would need it.

This is rewarding example to consider. It is a natural case in which to invoke Grice’s Maxim of Quantity, and applications of this maxim in general show his theory to best advantage. However, if we look for explanatory detail, the results are far from satisfactory.

Invoking Grice’s Maxim of Quantity, you could account for this implicature by imagining my wife asking herself what the point could be of my making this utterance. We both believe (and believe that we believe, and so on) that I haven’t told her anything about needing the car this afternoon; it’s no news. So the point of what I said must have been to do something else. A search for the point leads to the hypothesis that I mean that I will need the car.

This explanation is really not very satisfactory: (1) it doesn’t explain how the interpretive reasoning gets started, and (2) it offers little or no hope of showing how it reached the conclusion that it does.

The problem with (1) has to do with belief. It isn’t enough to say that my utterance is trivial when taken literally, since it is not in general true that we look for a hidden meaning when the literal message is trivial. The reason for this is that the contents of the conversational record may not coincide with what is mutually believed. Take Stalnaker’s barber shop example, in which the barber offers you the following piece of news: “We’ve been having nice weather.” Both of you *know* that the weather has been nice, and you both know that you know it. But you don’t look for a hidden message: the barber means that the weather has been nice, and that is about all that he means. The problem is, why would it be unreasonable for my wife to interpret my remark about the car as like the barber’s remark—an attempt to make conversation? Furthermore, there are many cases other than ones involving trivial assertions in which we do reconstruct implicatures; and we need some general way of explaining why this occurs.

Here it seems to me that computational work on planning provides a much more promising approach. In almost any sort of conversation we always feel compelled to reconstruct the plans of our interlocutors; and we want to see how the message we ascribe to them fits into a model of their purposes. If we can’t do this, we are likely to resort to accommodation in



order to make it fit. Triviality is one feature that makes it more difficult to achieve this sort of fit; and this is why an assertion that is literally trivial is likely to give rise to implicatures.

As for (2), I'll discuss the Maxims in more detail later, in section 13. But it should be clear that the Maxim of Quantity, and Grice's other maxims, really provide no clues as to how my wife could arrive at the proposition that I need the car. Presumably there are many relevant, informative, true things I could have said at this point in the conversation; and much stronger constraints are needed to limit the search initially and guide it to the one correct target, to the proposition that I'll need the car.

Here too, it seems to me that reconstructed plans are helpful. This conversational implicature seems to me to be induced by the recognition of the following plan. [I am reviewing my plans for the afternoon.] (a) I will need the car this afternoon; but (b) I can't assume that I can use it without my wife's agreement; and (c) the first step in negotiating for her agreement would be to tell her that I will need the car; but (d) I haven't yet told her that. So, since she is within conversational range (let us assume), it is natural for me to establish the immediate practical goal of telling her that I will need the car, with the expectation that this will enable her to recognize my domain plan.

But there are many ways of telling her this, and at this point I need to engage in discourse planning to choose a way of asserting that I'll need the car. The question is this: why is saying "I didn't tell you I'll need the car this afternoon" a feasible way of achieving my discourse goal?

The success of this discourse strategy, I think, depends in part on the fact that the sentence is negative and comes at the beginning of a discourse unit. Such sentences invite the hypothesis that they express a *lack*, and in fact point to an obstacle in a plan (of the speaker's, or perhaps someone else). This means that it is a default rule in interpreting discourse to seek such an interpretation of a negative sentence in such a position. For instance, if I begin a conversational unit by saying "There isn't a doorstop in this room," it would be appropriate for you to say "Why would you want a doorstop?" and it would be disingenuous of me to say "I didn't say I wanted a doorstop." Thus, I can suppose that saying "I didn't tell you I'll need the car this afternoon" will launch a search on my wife's part for a plan of mine that would be thwarted by my not telling her that I needed the car. If she has a normal ability to recognize domain plans, this should suffice to meet my discourse goal.

As for why I should choose this way of saying it—rather, say, than "I'll need the car this afternoon"—the more complicated sentence is likely to be salient in cases in which I had intended to speak to her about my plans but had forgotten. Here I have detected an obstacle in my own plans, and that obstacle is foremost in my mind. And it is just in such cases that I am

more likely to use a phrasing like "I didn't tell you I'll need the car this afternoon."

To a process of monitoring our interlocutors' plans and asking how a postulated meaning agrees with these plans, I believe we should also add a *mutual plan of the conversation*. In discussing the idea of the conversational record, I alluded to constraints on the way in which this record develops. Discourse must have a structure that enables us to anticipate, reconstruct, and remember the progress of a conversation. Many implicatures are generated by such *discourse expectations*, such as the implicature in the following dialog that Harry is home.<sup>31</sup>

[Two people, A and B, are talking at work about a fellow worker.]

A: "Where's Harry?"

B: "He's sick."

This implicature, I think, is set in motion by the discourse expectation that the normal response to a question should help to answer the question.

### 13 Grice's Account of Implicature

In this section I provide further background concerning Grice's account of conversational implicature, and elaborate on the criticisms I referred to earlier in this paper.<sup>32</sup> This section is self-contained and can be omitted by readers who do not want to hear more about Grice.

Grice finds the source of conversational implicatures in the "cooperative principle." He does not explain this principle in much detail, merely saying that a participant in a conversation can always opt out, saying "My lips are sealed: I will say no more." In accounting for actual cases of conversational implicature, he appeals to four maxims: Quantity, Quality, Relation, and Manner. Grice's explanations of particular implicatures involve reconstructed bits of reasoning, showing how the implicature would be a reasonable thing to guess the speaker to mean, given the circumstances and what was literally said, or "said in the favored sense."

I believe that there is some truth in founding a central class of implicatures on the cooperative principle, and that requirements such as the *mutuality* of the conversational record may point to a deep level at which conversation must be cooperative in order to be possible at all. And accommodation, with its component process of goal sharing, is obviously cooperative. Nevertheless, not all conversational implicatures seem to be based on cooperation in any realistic sense; certainly, implicatures are possible in situations that can only be described as hostile and uncooperative. To apply the cooperative principle to such examples, you would have to appeal to shared goals that are relatively trivial, such as the goal of

speaking in English. These don't seem very promising as explanations of pragmatic phenomena.<sup>33</sup>

I believe that it helps us to get a better grip on the relevant sort of cooperation if we think not so much of shared *domain goals* as of a shared sense of where the conversation has been and of where it is heading: the common plan of the conversation. Sharing a plan of the conversation may involve shared goals, but these have to do with discourse rather than with the subject at hand, and such goals, along with a common sense of the conversational record, can be shared even though the participants have few domain goals in common. Of course, many conversations also involve shared goals having to do with the task at hand, and these can enter into implicatures as well.

It isn't clear how Grice's maxims derive from the cooperative principle. But the internal looseness of Grice's theory is less important than the explanatory looseness of the maxims, their shapelessness when confronted with the phenomena.

Quantity, for instance, raises as many problems as it settles. This maxim tells us to say as much as is required. But that is a bit like asking a friend what you should wear to a party and being told to dress appropriately.

We obviously can't say all that we know. In a real, interactive conversation we can't even say all that we feel is relevant and helpful; we have to make a selection (either a reasoned or a random selection), and to hope that things will get sorted out satisfactorily as the conversation progresses. For example, when I tell you that Otto lives in Cambridge, I may be in a position to tell you that his mother knows that he lives there. The latter is more informative. It may even be pertinent. But I may still not choose to say it. And when I simply say that Otto lives in Cambridge, I do not implicate that his mother doesn't know that he lives there.

The moral is that we can never appeal *merely* to Quantity in explaining why, for instance, 'I tried to cash your check' implicates 'I didn't cash your check'.

The maxim of Relevance is even worse. Just about any implicature can be explained by invoking relevance. But such explanations are idle without some way of separating what is relevant at a given stage of a conversation from what is not.

My suggestion, then, is that we should replace the maxims with a general theory of group planning of public data structures, with special application to conversational planning. This theory should then yield many principles, more specific than the maxims, and deriving from the idea that what a speaker means should fit in and cohere with the conversational plan, and with the reconstructed plans of our conversational partners. The example of 'I didn't tell you I'll need the car' was supposed to illustrate the advantages of this approach in a specific case. (Given what I've said about

speaker meaning, this is really just a theory of how implicatures are meant.) A general advantage of such an approach is that, just as plans have a complex structure (containing, among other things, subplans, with habits or automatic routines at the lowest level), we should expect to find a variety of explanatory principles, of varying particularity, that serve to regulate conversational planning.

#### 14 *A Bit beyond Foundations: The Case of Not Knowing*

We test foundations by trying to put something on them. What needs to be put on these foundations is a broad spectrum of studies showing how to give detailed explanations of why implicatures are generated in some cases, and why they are not in others. Since these matters are very sensitive to context, it is really necessary to simultaneously develop a theory of the context, including representations of the features that affect implicature, as well as an account of how implicature is dependent on context. This is a special case—but a very complex special case—of developing a theory of data structures and of the reasoning procedures that interact with these data structures.

The present paper doesn't show enough detail to show that the foundation I have argued for will support the weight of implicature. However, McCafferty 1987 provides an extended study that goes much farther in this direction. And I would like to conclude by discussing one more example: the implicature that accompanies denials that someone knows something. This is a difficult case, one that has concerned me for a long time. I'm still not certain that I have gotten to the bottom of it.

Ordinarily, in saying that Arnold doesn't know that his shoelace is untied, I implicate that his shoelace is untied. The implicature appears to be conversational. For one thing, the implicature disappears in contexts in which it would be natural to stress the word 'know' (and in other contexts as well).<sup>34</sup> For another, 'know' is one of a class of predicates taking 'that' complements that carry an implicature of this sort: some of these, like 'aware', 'realize', and 'see', are synonyms or near-synonyms of 'know'. Others, like 'discover', 'find out', and 'remember', are more distant.<sup>35</sup>

I have changed my mind often about this implicature. One explanation that has survived many of these changes appeals to two things. (1) Often, whoever is speaking about a topic is liable to be given "expert" status. The expert is presumed to know the answers to the range of questions associated with the topic. Of course, this presumption is defeasible. But it has a certain *prima facie* force. Probably all of us who have unguardedly introduced a topic, only to be pressed with embarrassing questions, have suffered from this rule. (The worst case is to read an academic paper.) In a

relatively nonstructured situation, the role of expert can shift: it usually falls on the last person who has made an undefended assertion about a topic.

This has the consequence that in the absence of anything suggesting the contrary, there will be a presumption that the speaker who says "Arnold doesn't know that his shoelace is untied" will be presumed to know whether Arnold's shoelace is untied. (An added factor in this case is that the speaker doesn't choose to use the version with 'whether' rather than the version with 'that'. Because of the semantic differences between 'whether' and 'that', this would have the effect of hedging.)

ch (2) Denying that someone knows something will have a *point*. The question of whether someone knows something doesn't ordinarily arise unless we are concerned to praise or blame him, or to explain or predict some action, opinion, or the like. But, on the supposition that a proposition is false, denying that someone knows the proposition can serve neither of these purposes. It is no defect to be ignorant of what is false. And if something is false, the connections are undercut between the hypothesis that someone doesn't know it and his actions. (This is why, when a proposition is in doubt at a stage of a conversation,<sup>36</sup> and we are interested in someone's actions or opinions, we speak of whether he *believes* the proposition, or *knows whether* it is true, not of whether he knows it.)

Of course, there is another explanation, like the one that I invoked in a previous example: negative sentences like this one suggest obstacles to plans. That would have to be relaxed a bit, since in this case it is more an obstacle to Arnold's general welfare that is at stake. The explanation depends on the fact that Arnold's not knowing *p* couldn't be an obstacle to anything unless *p* were true. This explanation works over a wide range of examples, but I suspect that it would be hard to make it do justice to the full range of cases.

The explanation would have to be filled in with a number of rather particular, perhaps even listlike, principles about conversation and reasoning. And the trouble with lists is that it is very hard to tell when we should call them complete. But perhaps cognitive psychology could provide some help in organizing the lists. Also, the reasoning I have been describing is highly defeasible. To make sense of it, we need a theory of default reasoning. But we badly need such a theory anyway: in logic, cognitive psychology, and artificial intelligence. My hope is that these disciplines can be made to work together in accounting for conversational implicatures.

I'll conclude with a brief and very incomplete attempt to explain why stressing the word 'know' cancels the implicature in 'Arnold doesn't know that his shoelace is untied'.

I favor a pragmatic account of contrastive stress: I believe that stress conventionally implicates that a list of alternatives of a certain form is open at a certain stage of the conversational record. I think of such a list as a pair

$\langle \lambda x A, Y \rangle$ , where *Y* is a (possibly empty) set of values of the variable '*x*'. For instance, the alternative list  $\langle \lambda x(\text{Arnold kissed } x), \{\text{Betty, Sue}\} \rangle$  might be a reasonable background for 'Arnold kissed Sue', with heavy stress on 'Sue'. We must think of sets of alternatives as part of the conversational record. (It clearly would not do to identify the sort of *active* opening of options that goes along with a set of alternatives with a mere absence of suppositions. More is conveyed by contrastive stress of 'Sue' in 'Arnold kissed Sue' than absence of the supposition that Arnold didn't kiss someone else.)

It follows from what we have said about rules of conversation that contrastive stress can be used as a way of *implicating* that certain alternatives are open. Which alternative list is called up by such an implicature is heavily dependent on the context.

Second, stressed negatives are usually used, when they are used assertively, to exclude certain members of a list of alternatives and to force the remaining alternatives. The following case is typical. "She's in her office or at home. But she isn't in her office. So she must be at home." (There also often is implicature, which seems to be conversational, to the effect that the excluded alternative was the favored one, so that it is somewhat surprising that it is excluded.)

Third, if an alternative is presented to knowledge, we expect it to be belief. In fact, the alternative set  $\{\text{know, believe}\}$  is so natural that it is hard to force a different alternative set involving knowledge.

From these things, we can infer that 'Arnold doesn't know that his shoelace is untied' will ordinarily be used to say that Arnold believes that Arnold's shoelace is untied, though he does not know it. Ever since Plato, philosophers have known that knowledge is not the same thing as true belief. Even so, it remains true that we do not ordinarily claim that someone believes something without knowing it unless we want to cast some doubt on the belief: to open the alternative that it is false. Thus, any implicature that the belief is true will be cancelled in this case.

One useful moral of this case study is that we should not think of the conversational record merely as an accumulation of the suppositions that are in force at a stage of the conversation. It is a much richer and more diverse structure than this; and it contains, among other things, representations of the alternatives that are conceived as open. It is instructive, too, that this is just what is needed in order to represent interrogative speech acts.

It is perhaps a more important moral that considering evidence having to do with contrastive stress led us to a small, closed list of notions: the list consisting of knowledge and belief. And earlier, we were led to the thought that in order to obtain satisfactory explanations of implicature we would need to invoke such lists, and that it would be nice to have tests for their completeness.

One fruitful source of evidence bearing on these lists, I would like to suggest, is contrastive stress.

#### Notes

1. A good example is Russell's theory of descriptions, according to which 'the' generates both an existential claim and a uniqueness claim. An enormous philosophical literature developed over the existential claim and relatively little attention was paid to uniqueness, though there are obvious counterexamples, and the theoretical consequences of doing justice to them are probably more far-reaching. See McCawley 1985.
2. The paper has been reprinted so often that it is not very useful to cite page numbers. The passage I refer to begins with the fourth paragraph of part III.
3. Austin 1962a. This book consists of manuscript notes for lectures given regularly at Oxford since 1947, and revised after Austin's death by G. Warnock.
4. See Grice 1975 and Grice 1978 for published excerpts of the theory. The most extensive presentation of the theory is in Grice's William James Lectures, which have been widely circulated in unpublished form.
5. I myself advocated a semantic rather than a pragmatic approach to category mistakes in Thomason 1972. I have never been entirely happy with semantic theories of discourse anomaly even in this simple case, but I remain sympathetic to anyone who prefers these theories to pragmatic approaches because of their rigor, and for lack of anything better to take their place.
6. See, for instance, Kempson 1975, which makes points very similar to those I am summarizing here, with supporting evidence and many references to the literature.
7. This is pretty much the same as Stalnaker's proposal in Stalnaker 1972.
8. I would also include speaker meaning, for example, meaning that it's getting late, or meaning for someone to open the window, as a kind of act-scheme, if not an act.
9. The closest thing I have been able to find in the literature to a formulation of this rule is in Heim 1983: "For every definite, update an old card." This presupposes that a reference must already be established for every definite noun phrase; in effect, Heim's rule is the special application of our general rule to deictic uses of the definite article.
10. Here is what one pedant (Burman (1975, 79)) has to say: "This is one of the most fatuous of proverbs, as usually applied. Quite clearly the exception cannot prove a rule; quite the opposite; it *disproves* it."
11. Where there is competition and conflict, there is less likely to be accommodation; so in institutions like games, where bickering is undesirable, we are likely to regiment away the need for accommodation by codifying a more or less elaborate body of explicit rules.
12. See Kamp 1981, Heim 1982, and Heim 1983.
13. See Fillmore 1975.
14. For instance, in Case (1) you believe that I believe that you intend to find a service station; in Case (2) you probably don't believe this.
15. Work such as that of Litman and Allen seems to be moving in this direction; see chapter 17 of this volume.
16. Or match well enough. In many realistic situations an approximation of coordination will suffice, especially if there is a mechanism for detecting and correcting miscoordinations. (Such mechanisms clearly exist in human communication; much of the redundancy in conversation serves to check coordination. For instance, this seems to be why it is natural not only to ask a question but to explain the point of the question, as in an example cited by Pollack in chapter 5: "I want to talk to Kathy, so I need to find out the phone number of St. Eligius.")
17. The job is impossible if these assumptions are relaxed; see Halpern and Moses 1984.
18. But obviously, the distinction is needed in some computational applications: cases in which a variety of users can access a database, and need to be allowed different privileges in updating the database. If we had a system of this sort that also had sophisticated discourse capabilities, we might well want to separate presumption from belief in order to have effective conversational capabilities as well as integrity of data. Such a system, like a human being, would have to be capable of harboring multiple systems of suppositions.
19. McCafferty 1987 is an exception.
20. In a way, this point is trivial, since single intentions can count as (degenerate) plans. In fact, however, we hardly ever recognize such simple plans in isolation.
21. This would require a distinction between "first-order requirements"—what is required by conversational rules other than the rules of accommodation—and the other rules. Also, the "such-and-such" clause would have to be either replaced with explicit conditions, or eliminated by invoking default reasoning.
22. Like the version of the definition that I presented in section 8.1, this is the declarative case only of the definition. But this version involves a self-referential intention, rather than two intentions, one of which refers to the other. I prefer this second version of the definition, though I think it would be very difficult to argue on the basis of linguistic evidence for one version or the other.
23. At least on an indefinite construal of speaker meaning; and this construal is clearly the target of Grice's analysis. (I realize that this remark is obscure; I'll clarify it soon.)
24. I am not denying that the reflections and convolutions of Grice's definition are part of the subject matter; but the idea is to locate them elsewhere. In particular, on the approach that I am trying to develop they are concealed in the mutuality of the conversational record.
25. This is plausible at least for human hearers. It would not apply, of course, to natural-language understanding systems that depended only on parsing.
26. Of course, this is a secondary meaning. The primary meaning—*what* A meant by saying "There's a gas station around the corner"—is that there's a gas station around the corner.
27. Or at least, it creates the meaning that A doesn't suspect that the gas station may be closed. If we add the presumption that A is a local expert, we get the proposition that the station is open.
28. It is clear enough from Grice's tests for distinguishing types of implicature that conventional implicatures will have something to do with linguistic convention. It is not so clear that they are always implicatures. Suppose that you say "Billie likes ice cream" and I say "Susie likes ice cream too." According to Grice's terminology, I "conventionally implicate" that someone else likes ice cream. But do I *implicate* it in this case?
29. There are, I think, some implicatures that can't be disavowed this way; in particular, implicatures that are entailed or presupposed by what is said.
30. New foundations for pragmatics should provide a theory-based classification of implicatures. I hope that this is at least a first step toward such a classification.
31. I'm referring to an implicature that, like most implicatures, involves a certain amount of stage setting—more than I've provided in this case. Accommodate me.
32. Though I use the word 'criticisms', I don't mean to suggest that I think of myself as a party to any significant philosophical dispute with Grice. I have the impression that Grice was tentative in the William James Lectures about his account of conversational implicatures, and may not have been entirely satisfied with its ability to generate detailed explanations of a wide range of cases. On the whole, I believe that linguists and computer scientists have taken the details of Grice's theory more seriously than they

perhaps should have. It is important to remember that Grice's William James Lectures were never prepared for publication. Since I haven't corresponded with Grice about the ideas in this paper, I don't know how seriously he takes the tactics of his theory, or what the actual points of agreement and disagreement are between our current views.

33. Of course, the cooperative principle is useful (though at a general level that needs to be supplemented with more specific rules) in explaining some things that are related to conversational implicature, such as the inference that the speaker doesn't know whether the coffee beans are in the cupboard or the refrigerator if he says that the coffee beans are in the cupboard or in the refrigerator. But at this point, the inferences that are made seem to be shading off into natural meaning, and I am not sure that I want to call them implicatures; compare the inference that the speaker believes what he asserts.
34. This is an application of Grice's cancellability test; it assumes that contrastive stress does not affect literal meaning, an assumption that I believe can be justified.
35. In this paragraph, I have applied Grice's cancellability and nondetachability tests for conversational implicature.
36. I hope it is clear that I am invoking two sorts of plans in accounting for implicatures: a common plan of the conversation, and plans of the participants. (The latter plans may be reconstructed by other participants in the conversation.)

### References

- Allen, J. (1983). Recognizing intentions from natural language utterances. In M. Brady and R. Berwick, eds., *Computational models of discourse*, pp. 107–166. Cambridge, MA: MIT Press.
- Allen, J., and C. R. Perrault (1980). Analyzing intention in dialogues. *Artificial Intelligence* 15, 143–178.
- Anderson, A., and N. Belnap, Jr. (1975). *Entailment*, vol. 1. Princeton, NJ: Princeton University Press.
- Austin, J. (1962a). *Sense and sensibilia*. Oxford: Oxford University Press.
- Austin, J. (1962b). *How to do things with words*. Oxford: Oxford University Press.
- Brown, G., and G. Yule (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Burman, T. (1975). *The dictionary of misinformation*. New York: Thomas Y. Crowell Co.
- Carnap, R. (1947). *Meaning and necessity*. Chicago: University of Chicago Press.
- Cohen, P., and H. Levesque (1980). Speech acts and the recognition of shared plans. In *Proceedings of the Third Biennial Conference, Canadian Society for Computational Studies of Intelligence*, pp. 263–271. Victoria, British Columbia: Canadian Society for Computational Studies of Intelligence.
- Cohen, P., and H. Levesque (1990a). Persistence, intention, and commitment. This volume.
- Cohen, P., and H. Levesque (1990b). Rational interaction as the basis for communication. This volume.
- Fillmore, C. (1975). *Santa Cruz lectures on deixis*. Bloomington, IN: Indiana University Linguistics Club.
- Frege, G. (1960). On sense and reference. In P. T. Geach and M. Black, eds., *Translations from the philosophical writings of Gottlob Frege*, pp. 56–78. Oxford: Basil Blackwell.
- Grice, H. P. (1957). Meaning. *The Philosophical Review* 66, 377–388.
- Grice, H. P. (1961). The causal theory of perception. *Proceedings of the Aristotelian Society*, supplementary vol. 35, 121–152.
- Grice, H. P. (1975). Logic and conversation. In P. Cole and J. Morgan, eds., *Syntax and semantics 3: Speech acts*, pp. 41–58. New York: Academic Press. Also in D. Davidson and G. Harman, eds., (1975). *The logic of grammar*, pp. 64–75. Encino, CA: Dickenson Publishing Co.
- Grice, H. P. (1978). Further notes on logic and conversation. In P. Cole, ed., *Syntax and semantics 9: Pragmatics*, pp. 113–128. New York: Academic Press.
- Grosz, B., and C. Sidner (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics* 12, 175–204.
- Grosz, B., and C. Sidner (1990). Plans for discourse. This volume.
- Halpern, J., and Y. Moses (1984). *Knowledge and common knowledge in a distributed environment*. Technical Report, IBM.
- Heim, I. (1982). *The semantics of definite and indefinite noun phrases*. Amherst, MA: Graduate Linguistics Student Association, University of Massachusetts.
- Heim, I. (1983). File change semantics and the familiarity theory of definiteness. In R. Bäuerle et al., eds., *Meaning, use, and the interpretation of language*, pp. 164–189. Berlin: Walter de Gruyter.
- Kamp, J. A. W. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, and M. Stokhof, eds., *Formal methods in the study of language*, pp. 277–322. Amsterdam: Mathematical Center.
- Karttunen, L. (1973). Presuppositions of compound sentences. *Linguistic Inquiry* 4, 169–271.
- Kautz, H. (1990). A circumscriptive theory of plan recognition. This volume.
- Kempson, R. (1975). *Presupposition and the delimitation of semantics*. Cambridge: Cambridge University Press.
- Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.
- Lewis, D. (1979). Scorekeeping in a language game. *Journal of Philosophical Logic* 8, 339–359.
- Litman, D., and J. Allen (1990). Discourse processing and commonsense plans. This volume.
- McCafferty, A. (1987). *Reasoning about implicature*. Doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.
- McCawley, J. (1985). Actions and events despite Bertrand Russell. In E. LePore and B. McLaughlin, eds., *Actions and events: Perspectives on the philosophy of Donald Davidson*, pp. 177–192. Oxford: Basil Blackwell.
- Perrault, C. R. (1990). An application of default logic to speech act theory. This volume.
- Perrault, C. R., and P. Cohen (1980). A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics* 6, 167–182.
- Pierrehumbert, J., and J. Hirschberg (1990). The meaning of intonational contours in the interpretation of discourse. This volume.
- Pollack, M. (1990). Plans as complex mental attitudes. This volume.
- Russell, B. (1970). Mr. Strawson on referring. *Mind* 66, 385–395.
- Stalnaker, R. (1972). Pragmatics. In D. Davidson and G. Harman eds., *Semantics of natural language*, pp. 380–397. Dordrecht: D. Reidel.
- Stalnaker, R. (1973). Presuppositions. *Journal of Philosophical Logic* 2, 447–457.
- Stalnaker, R. (1975a). Pragmatic presuppositions. In M. Munitz and P. Unger, eds., *Semantics and philosophy*, pp. 197–213. New York: Academic Press.
- Stalnaker, R. (1975b). Indicative conditionals. *Philosophia* 5, 269–286.
- Stalnaker, R. (1978). Assertion. In P. Cole, ed., *Syntax and semantics 9: pragmatics*, pp. 315–332. New York: Academic Press.
- Strawson, P. (1950). On referring. *Mind* 59, 320–344.
- Thomason, R. (1972). A semantic theory of sortal incorrectness. *Journal of Philosophical Logic* 2, 209–258.
- Thomason, R. (1987). The multiplicity of belief and desire. In M. Georgeff and A. Lansky, eds., *Reasoning about actions and plans: Proceedings of the 1986 workshop*, pp. 341–360. Los Altos, CA: Morgan Kaufmann.