

The Grice is Right: Grice's Non-Cooperation Problem and the Structure of Conversation

Sam Berstler
Massachusetts Institute of Technology
Prepared for *Philosophical Perspectives*

Abstract. H.P. Grice seemed to rest his theory of conversational implicature on the assumption that speakers aim to cooperatively exchange information with each other. In the real world, speakers often don't. Does one of the most influential theories in twentieth-century philosophy of language rest on a mistake? Yes—but not in the way that philosophers have thought. I argue that Grice should have rested his theory on a different assumption: that speakers aim to *appear* to aim to cooperatively exchange information with each other. This proposal dissolves Grice's Non-Cooperation Problem but preserves Grice's central insights about the nature of conversational implicatures. More generally, it enables the Gricean to illuminate the structure of many non-cooperative or otherwise “non-ideal” conversations.

Keywords: H.P. Grice; conversational implicature; strategic speech; non-cooperative conversations; conversational genre

I. Introduction

In a bar downtown on a Friday night, Elena turns to her friend Arthur and asks:

Bar

Elena: 1. What do you think of this place?
Arthur: 2a. I'm getting a headache.

Elena infers that Arthur is getting a headache because of the thumping music and that consequently he doesn't like the bar. Elena doesn't have some special insight into Arthur's psychology. She's not a master mind-reader. In one of the most ordinary forms of communication in the world, Arthur simply meant more than what he said.

How did Elena know¹ that Arthur meant (2b) and not (2c) or (2d)?

2b. I don't like the bar because of the loud music, which is causing my headache.
2c. I'm worried that this headache is a symptom of brain cancer.

¹ My use of the word “know” is opinionated. In “Meaning” Grice proposed that a speaker who means *p* constitutively intends for her addressee to *recognize* that she means that *p* (Grice 1957). I am interpreting “recognizing” as “knowing.” For the purposes of this paper, nothing meaningfully hangs on this decision.

2d. You should immediately go to CVS and buy me Tylenol.

In “Logic and Conversation,” H.P. Grice famously answered: Elena relied upon her knowledge that she and Arthur are cooperating (Grice 1989 [1967]: 22-40). Grice’s resulting theory of *conversational implicature* became one of the most influential ideas in twentieth-century philosophy of language. Arguably, it laid the foundation for much of contemporary pragmatics as we find it.

It is therefore surprising that Grice’s theory seems to rest on such an obvious mistake.

The problem is right there at the beginning. In order to explain how speakers rationally make and interpret conversational implicatures, Grice assumes that they are conforming to the Cooperative Principle:

Cooperative Principle

Make your contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

(Grice 1989 [1967]: 26)

Obviously enough, speakers often don’t seem to conform to this principle. In the real world, speakers purposefully lie, meander, obfuscate, deflect, badger each other, and otherwise engage in not-very-nice behavior. And as we will see, in many of these conversations, speakers non-problematically make and interpret conversational implicatures. Checkmate, Grice?

Grice’s Non-Cooperation Problem (as I’ll call it) is an old and persistent objection to Grice’s theory of conversational implicature. It appeared first in the 1970s and has periodically resurfaced ever since.² But these days, the problem has a special urgency. More and more philosophers of language are investigating the structure of non-ideal conversations and the strategies that speakers employ within them. And in part because of this, more and more philosophers of language are critically assessing the simplifying assumptions that have guided much of the field.³ With its blithe confidence that speakers cooperate with each other and its jaw-dropping level of influence, Grice’s “Logic and Conversation” is a natural target for this kind of critique.

My aim here is to defend the Gricean from these critiques—not by rejecting the critiques themselves but by re-imagining Grice’s project in light of them. I agree that Grice made a mistake: he should *not* have rested his theory of implicature on the assumption that speakers aims to appear to aim to cooperatively exchange information with each other. But he *should* have rested it on a nearby assumption: that speakers aim to *appear*

² Versions of the problem are discussed in Keenan (1976), Harnish (1991 [1977]: 357 fn29), Bird (1979), Holdcraft (1979), Kasher (1982), Grice (1989: 368-370), Davis (1998), Thomason (1990: 355-357), Green (1995), Asher (2012), Asher and Lascarides (2013), Stokke (2016), Capellen and Dever (2021), Ellefson (2021), Keiser (2022), and Beaver and Stanley (2023).

³ Keiser (2022) and Beaver and Stanley (2023) are book-length treatments that specifically criticize idealization in the philosophy of language. Very recent articles that espouse similar kinds of critiques include: on non-ideal philosophy of language in general, Mühlbach (2022), Hesni (2024), Mühlbach and Ernst (2024), Engelhardt and Moran (forth.); on social meaning, Michaelson (2024), Nowak (forth.); on discourse structure, Camp (2018), Berstler (forth. b); on refusal, Kukla (forth.). Since at least the 1990s, feminist philosophers of language have argued that attending to the role of power in communication can generate insights about foundational issues in the field. For work in this tradition, see Langton (1993), Hornsby (1995), Maitra (2009), Tirrell (2012), Kukla (2014), and McGowan (2019). Finally, for the absolutely voluminous on what slurs and hate speech may show about foundational issues in the philosophy of language, see Anderson and Barnes (2022).

to cooperatively exchange information with each other.⁴ On this reinterpretation, Grice is right that the Cooperative Principle explains how we make and interpret conversational implicatures. He is just wrong about how and why it does so.

On my re-imagining of “Logic and Conversation,” the Gricean becomes a friend, not a foe, to those philosophers who wish to study non-ideal conversations. She acquires the resources to cast genuine explanatory light on their structure. And she can all of this within an almost entirely non-technical framework, with only small changes to the original theory, and with relatively minimal assumptions about the structure of conversations.

Here’s how I’ve arranged this paper. In the next section, I introduce a minimal pair that seems to falsify Grice’s theory. I also clarify what Grice’s Non-Cooperation Problem isn’t. After that, I unfold my positive view. To do this, I first discuss a case of failed misleading. In this case, the speaker intends to deceive her interlocutor but fails to do so. Then, I introduce the common ground framework (Stalnaker 2002, 2014) and, with its tools, show how to fully generalize my solution. In the last section, I discuss three kinds of conversations that my analysis helps to illuminate. Of particular importance are what I call *ritual conversations*: conversations in which the participants jointly aim to merely appear to aim to exchange information.

II. The basic problem

a. The minimal pair

What is Grice’s Non-Cooperation Problem? To start, I will sketch a minimal pair that philosophers and linguists have recently puzzled over.

Suppose that some sketchy cryptocurrency mogul—let’s call him, randomly, Sam—is talking to his lawyer, Rose. In order to help Sam out of his various legal predicaments, Rose needs information about Sam’s actual finances. They have the following conversation:

Bank account 1

Rose:	1a. Have you ever had a bank account in Hong Kong?
Sam:	1b. My company had one there for about six months.

In saying (1b), Sam means more than what he says. He means the conjunction of (1b) and (1c). Rose recognizes that Sam means (1c). And because she trusts him, Rose infers (1c):

Sam:	1c. No, I never had a personal bank account in Hong Kong.
------	---

In Grice’s terminology, Sam has *conversationally implicated* (1c).

We can check that Sam did, in fact, mean (1c). First, Rose can non-problematically presuppose (1c):

Rose:	1d. Why didn’t <i>you</i> have one?
	1e. But that’s a good place for personal accounts.

If Rose does want to double-check that Sam meant (1c), she can. But to do so, she ought to presuppose that Sam meant (1c) and that she is only seeking clarification or confirmation of this:

⁴ This move is presaged in Grice’s own work. In a single paragraph in the “Retrospective Epilogue,” Grice writes that in some non-cooperative conversations, the speaker’s “common objectives are spurious, apparent rather than real” (Grice 1989: 370). I discuss this passage in the conclusion.

1f. Just to be clear, *you* didn't have one, right?
 1g. So the answer is "no," right?

How did Rose know what Sam meant? Grice answers: she knows that she and Sam are cooperatively pursuing a shared project. In particular, Rose knows that she and Sam cooperatively aim to pool their beliefs or knowledge about Sam's financial situation.

Because of that, Rose knows that Sam is conforming to the Cooperative Principle. And because she knows *that*, Rose knows that Sam is conforming to the Four Maxims:⁵

Four Maxims

Quality:	speak truly;
	do not say things for which you lack adequate evidence
Quantity:	say no more or less than is required
Relation:	be relevant
Manner:	be perspicuous

(Grice 1989 [1967]: 26-27)

Grice proposes that Rose tacitly reasons in something like the following way. First, Rose asks: did Sam just mean what he says? She concludes that he does not. If Sam had just meant what he said, then he would not have been cooperative. More precisely, Sam would not have conformed to the maxim of Relation.⁶ To reply in a *relevant* way to her question, Sam would have had to answer the question. If Sam just meant what he said, he would not have answered. But, Rose reasons, Sam *is* cooperative. So Sam must have conformed to Relation. So Rose searches for a reinterpretation of what Sam means—a reinterpretation on which Sam *does* answer the question.

Now consider the same dialogue fragment but in a slightly different context. Suppose now that Sam is not talking to his attorney but an investigative journalist, Yuna.⁷ And let us suppose that Yuna is asking about Sam's *Swiss* bank accounts. Sam has personal accounts in Zurich, which he is using in order to launder money. Yuna knows about these accounts, and she knows that Sam is using them to commit financial crimes. They have the following conversation:

Bank account 2

Yuna:	2a. Have you ever had a bank account in Switzerland?
Sam:	2b. My company had one there for about six months.

As before, in saying (2b), Sam intends for Yuna to infer (2c). And as before, Sam intends for Yuna to infer (2b) because he intends for Yuna to believe that he means (2b).

⁵ Grice is coy about the precise relationship between the Cooperative Principle and the Four Maxims. He just says that following "will, in general, yield results in accordance with the Cooperative Principle" (Grice 1989 [1967]: 26).

⁶ In their initial formulation of the problem, Asher and Lascarides diagnosis this as a Quantity violation. But that is, in part, because they use their own reconfigured versions of the maxims (Asher 2012, Asher and Lascarides 2013). In any case, whether this implicature is ultimately a Quantity-based or Relation-based implicature is dialectically irrelevant.

⁷ In almost all recent discussions of this case, philosophers and linguists cite a real-life exchange on cross examination between a prosecutor and a defendant (Asher 2012, Asher and Lascarides 2013, Keiser 2022, Beaver and Stanley 2023), attributed to a passage in Solan and Tiersma (2005), which itself draws on the trial of *Bronston vs. United States*. But court rooms are special contexts: official speech within courtrooms is rigidly regulated, and it's not always clear how such speech relates to its informal counterparts (Harris 2020: 15-16). To sidestep a dialectical issue, I have just updated the example. I discuss court room cases in IV.b.

Sam: 2c. No, I never had a personal bank account in Switzerland.

As before, Yuna can felicitously continue the conversation under the presupposition that Sam meant (2c):

Yuna: 2d. Why didn't *you* have one?
2e. But that's a good place for personal accounts.

And as before, Yuna can double-check that Sam meant (2c). To do this in a felicitous way, she must presuppose that Sam did mean (2c):

2f. Just to be clear, *you* didn't have one, right?
2g. So the answer is "no," right?

But now the Gricean has a problem. Once again, we can ask: how does Yuna know that Sam means (2c)? The Gricean cannot provide the same story that she provided when she explained Rose's implicature. For when Yuna tacitly asks herself, "Did Sam just mean what he said?" she has no reason to answer no. Yuna knows that she and Sam do *not* aim to cooperatively pool their beliefs about his finances. So Yuna has no reason to believe that whatever Sam means, he intends to answer her question. Yuna has no reason to search for a reinterpretation of Sam's speech act.

This is Grice's Non-Cooperation Problem. It seems to show that whatever Grice gets right about conversational implicature, his appeal to the Cooperative Principle is just plain wrong.

b. What about general non-cooperation in conversation?

Let me pause to explain what the problem is and isn't. Not every case in which interlocutors fail to conform to the Cooperative Principle is a problem for Grice. For the Gricean, what generates the problem is this assumption:

Key Assumption

Whenever a speaker successfully makes a conversational implicature, the addressee reasonably believes⁸ that the speaker is conforming to the Cooperative Principle

The Gricean could justify her Key Assumption with a strong claim about what interlocutors always *do*:

Universal Conformance

In any conversation, interlocutors always conform to the Cooperative Principle.

If interlocutors always *do* conform to the Cooperative Principle, then it's plausible that interlocutors nearly always reasonably believe that they are each conforming to the Cooperative Principle. Then the Gricean can just help herself to the Key Assumption.

But the Gricean *shouldn't* pursue this strategy. On any reasonable interpretation of the Cooperative Principle, **Universal Conformance** is too implausible to take seriously.

Consider this case:

⁸ I use "reasonably believe" instead of "know" so that it is not completely trivial to falsify the Key Assumption. If the Key Assumption requires an addressee to *know* that her interlocutor is conforming to the Cooperative Principle, then any case of successful, deceptive misleading will falsify the Assumption.

Kitchen Mayhem

Suppose that Lucy and Dan are cooking Thanksgiving dinner together, and they want to produce an excellent meal.

Dan: Where did you put the saltshaker?
 Lucy: It's behind you, you dipshit!
 Dan: Urgh, stop *yelling at me!*

Lucy flusters Dan so much that he spills too vinegar into his pan and wrecks his sauce.

What is the “accepted purpose or direction” of Lucy and Dan’s conversation as they cook? Something like: coordinating their actions and knowledge in order to produce an excellent Thanksgiving meal. Let us suppose that Lucy knew that in yelling at Dan, she would most likely fluster him and undermine his ability to execute on his tasks.

Then in saying, “You dipshit!” Lucy knowingly failed to “make [her] contribution as such is required, at the stage at which it occurs,” given their “accepted purpose or direction of the talk exchange.” She spoke in a way that she knew would *undermine* their shared project.

This sort of scenario is the most ordinary thing in the world. We regularly speak in ways that violate the Cooperative Principle: we lose our tempers, or we are not thinking clearly, or we have some other personal aim which conflicts with our shared conversational goal. We are angry, so, like Lucy, we violate Manner. We want to deceive our interlocutor and so violate Quality. We want to show off, drone on too long about metaphysics, and so violate Quantity. Universal Conformance simply can’t be right.⁹

But the failure of Universal Conformance does not falsify the Key Assumption. It only falsifies one *way* of motivating the Key Assumption. So any conversation that falsifies Universal Conformance is not necessarily an example of Grice’s Non-Cooperation Problem.

To generate a clear problem for the Gricean, we need a case like **Bank Account 2**. These cases are problematic because, within them, *two* facts are simultaneously true: (i) the speaker unambiguously successfully conversationally implicates something to an addressee and (ii) the addressee reasonably believes that the speaker is not conforming to the Cooperative Principle.

c. *Objection: Sam is cooperative.*

Isn’t there a sense in which Sam and Yuna *are* cooperating? Yes. But this fact does not meaningfully undermine the Non-Cooperative Problem. Let me explain.

Plausibly, two interlocutors are cooperating if they are rationally working together in order to achieve a joint goal. Do Sam and Yuna have any joint goals? Plausibly, they do. They jointly aim to understand each

⁹ This is different from Grice’s point that “collaboration in achieving exchange of information or the institution of decisions may coexist with a high degree of reserve, hostility, and chicanery” (Grice 1989: 369). Here Grice is pointing out that the *mere* fact that I have (say) called you a “dipshit” doesn’t obviously undermine our shared conversational goal. If it does not undermine our shared goal, then it is an example of interpersonal “hostility” but not a violation of the Cooperative Principle.

other.¹⁰ And they jointly aim to conduct a journalistic interview, in which Yuna takes the role of interviewer and Sam takes the role of interviewee.

Are they rationally working together to achieve those goals? Again, plausibly they are. Sam and Yuna are speaking in a mutually intelligible language; they are taking turns speaking, so that each person can hear the other; they are seated at a socially appropriate distance away from each other; Yuna has brought a voice recorder, which Sam helpfully speaks into. And so isn't it true that Sam is conforming to the Cooperative Principle and that the Key Assumption holds?

The problem is that insofar as Sam is conforming to the Cooperative Principle, this sort of cooperativity can't *explain* the sorts of facts that the Gricean needs to explain. Consider again the moment at which Yuna asks: "Did Sam just mean what he said?" The classical Gricean analysis depends upon the following five claims:

1. Yuna reasonably believes that (i) Sam did not only mean what he said.
2. Yuna reasonably believes that (ii) whatever Sam meant, he answered her question.
3. Yuna reasonably believes that (iii) she and Sam are cooperatively pursuing some goal.
4. Yuna reasonably believes (i) *because of* her belief in (ii).
5. Yuna reasonably believes (ii) *because of* her belief in (iii).

But from the fact that (iii) holds, Yuna cannot necessarily infer (ii) and thus she cannot necessarily infer (i). Just because Sam and Yuna are working together to understand each other and to conduct an interview, it does not follow that Sam is rationally required to directly answer her question. In other words, in order to explain why Yuna moves from (iii) to (ii) and then to (i), we need to assume a further claim:

6. Yuna reasonably believes that (vi) in virtue of the nature of Sam and Yuna's joint goal and the fact that Sam knows the answer to Yuna's question, Sam is rationally required to answer Yuna's question.

If Sam and Yuna intended to *pool their beliefs or knowledge*, then Yuna could reasonably believe (vi), and the Gricean wouldn't have a problem. But that's exactly the claim that we are denying. Yuna intends to extract information from Sam that Sam intends to conceal, and Yuna knows this.

To be more precise, we can reformulate the Key Assumption:

Key Assumption 2

Whenever a speaker successfully makes a conversational implicature, the addressee reasonably believes that if the speaker had not made that conversational implicature, she would have violated the Cooperative Principle.

If Sam had not meant (2c), he would not have violated the Cooperative Principle. Given the nature of Sam and Yuna's actual shared project, Sam is not rationally required to directly answer Yuna's question. So to fail to answer is *not* to fail to "make his contribution such as is required, at the stage at which it occurs, given the accepted purpose or direction of their talk exchange."

¹⁰ For Grice, each interlocutor constitutively intends for the other to recognize what she means (Grice 1957, 1989). This doesn't yet show that the interlocutors are *cooperating* with each other, since this doesn't yet show that they have a *joint* project. I think it's the turn-taking behavior that shows that they have a joint goal to make themselves intelligible to the other. But nothing much hangs on this.

d. Objection: This is the wrong way of interpreting the Cooperative Principle.

Does this argument hang on how we interpret or ground the Cooperative Principle and the Four Maxims? Not really. I've been assuming a particular understanding of the Cooperative Principle. In the interest of transparency, I will make that reading explicit, and I will contrast it with another, nearby one. But then I will show that the argument doesn't *depend* upon my interpretation either.

Here's my preferred, minimalist interpretation of the Cooperative Principle. The Cooperative Principle is just a specific instance of a principle of instrumental rationality. That principle says: take the means to pursue your ends. From that principle, we can derive a second one: take the means to pursue your *joint* ends. And from that principle, we can derive a third principle: take the means to pursue whatever joint ends you have within your talk exchange. This third principle is what I think the Cooperative Principle is.

On this interpretation, the Cooperative Principle and, derivatively, the Four Maxims are grounded within our rationality. They govern us insofar as we are rational beings pursuing a joint project.¹¹

This version of the Cooperative Principle does *not* enjoin us to cooperate in one sense of the phrase. It does not itself require us to *actually* adopt any particular joint goal. Compare the Cooperative Principle with what Jesus (or morality) commands. When Jesus commands, "Cooperate with each other!" he does not mean, "Pursue whatever joint ends you already have in a rational way!" He is commanding us to actually *undertake* joint projects with each other.

But suppose instead that we interpret the Cooperative Principle or the Four Maxims or both in some other way. Suppose that we interpret the Cooperative Principle and the Four Maxims as *social norms* or even *constitutive norms*, which govern conversation in virtue of the kind of activity it is.¹² We can even interpret the Cooperative Principle as *requiring* us to identify with each other's conversational aims, not just to rationally pursue whatever ends we antecedently jointly share.

On this interpretation, we have reason to conform to the Four Maxims insofar as we have reason to conform to these norms. And in any given conversation, these norms can provide us with additional *motivation* for cooperating with each other. If you ask me, "Where is the train station?" I may truthfully answer you, just because I value conforming to the social norm that requires me to adopt your end.

¹¹ This interpretation seems to best conform to the spirit of Grice's writing. For example, Grice says that his aim in "Logic and Conversation" is "to see talking as a special case or variety of purposive, indeed rational behavior" (Grice 1989 [1967]: 28). In the "Retrospective Epilogue," he repeats that "it is the rationality or irrationality of conversational conduct which I have been concerned to track down" (Grice 1989: 369). Daniel W. Harris, who also subscribes to this interpretation of the Cooperative Principle, has labeled it "Paleo-Griceanism" (Harris p.c.). One advantage of Paleo-Griceanism is that it interacts in predictable ways with our theory of joint action. If we deploy a richer theory of joint planning and agency (Thomasson 1990, Bratman 2014), we can straightforwardly enrich Griceanism.

¹² There *is* some textual evidence that Grice had considered and favored such an interpretation. In "Logic and Conversation," Grice writes that he would "like to be able to show that observance of the Cooperative Principle and [M]axims is reasonable (rational) along the following lines: that anyone who cares about the goals that are central to conversation/communication (such as giving and receiving information, influencing and being influenced by others) must be expected to have an interest, given suitable circumstances, in participating in talk exchanges that will be profitable only on the assumption that they are conducted in general accordance with the Cooperative Principle and the [M]axims" (Grice 1989 [1967]: 29-30). It is possible to read Grice as proposing the following: insofar we *generally* care about communication, then in any conversation, we *always* have at least some reason to conform to a *social practice* that properly functions to facilitate communication. But given Grice's cageyness, it's hard to know what to make of this passage. Stotts (2024) develops the view that the Four Maxims are social norms, although the Non-Cooperative Problem does not motivate her analysis.

It is a deep question which view the Gricean ought to adopt and what (if anything) turns on this question. But we can't and needn't embark on that theoretical journey now. When it comes to the Non-Cooperation Problem, shifting to this second interpretation of the Cooperative Principle does not gain us any meaningful ground.

To see why, consider once again Yuna's interpretative process. On this view, when Yuna asks, "Did Sam just mean what he said?" she should ask herself, "Is Sam conforming to the social norm of Relation?" In other words, she should ask, "Is Sam conforming to the social norm that requires him to directly answer my question?" Given what she knows about Sam's actual aims, she is not in a position to answer yes. (Similarly, if she asks herself, "Is Sam conforming to the norms that require him to tell me the truth?" she is not in a position to answer yes.) And so Yuna is not in a position to reinterpret Sam's utterance. For all she knows, Sam has declined to answer her question. But this is the wrong prediction. Earlier, we had said that Yuna *knows* that Sam implicated (2c). That is the *datum* the Gricean wants to vindicate.

For the purposes of this paper, I don't think that the social norm interpretation is doomed—at least, not any more doomed than the interpretation of the Cooperative Principle that I prefer. The solution that I provide works for both.

III. The mere appearance of cooperation

Where exactly did Grice go wrong?

Let me try to channel the "non-ideal philosophy of language" zeitgeist for a moment. For these philosophers of language, the moral is clear: Grice should not have made conversational cooperation so central to his account. Insofar as Grice relied on the assumption that speakers are cooperatively exchanging information, his analysis reflects a widespread parochialism in the philosophy of language, one that abjures the messiness of ordinary talk and thereby distorts its subject matter beyond recognition. If there is anything worth salvaging in Grice's analysis, we will need to drop any appeal to cooperation in our theory.

As natural as this conclusion may seem, I am now going to argue that it is too quick. When a speaker successfully makes a conversational implicature in a conversation, I will argue, it is because she at least *aims to appear* to cooperatively pool information with her interlocutor. Sometimes, she *merely* aims to appear to cooperate with her interlocutor.

Crucially, what it is to appear to cooperate towards some goal systematically depends what it is to actually cooperate towards it. So even when a speaker merely aims to appear to cooperatively pursue some goal, her behavior will bear a systematic relationship to what she would do, were she *actually* to cooperatively pursue it. So cooperation remains central to my story of conversational implicature. So does the Cooperative Principle. They just don't always enter into the story in the way that Grice imagined.

To get an initial taste of what I am talking about, consider this actual interview between two non-cooperating figures. In this exchange, the *New York Times'* Andrew Ross Sorkin questions the disgraced cryptocurrency billionaire Sam Bankman-Fried. Here's how it opens:

Sorkin: Sam, I want to thank you for joining us this afternoon. I appreciate your willingness to have this conversation. As I said at the outset of today, there are a lot of questions that need to be asked and also need to be answered. As you know, a lot of people have been hurt, genuinely hurt. And my hope is that over the time we have together we can have a candid conversation about what happened—how it

happened. There are people who are angry, and they are seeking answers....I just want to say that I think our job as journalists is to have those conversations—is to ask those questions and seek those answers on behalf of the public.

(Sorkin 2022)

Throughout the interview, Bankman-Fried attempts to present himself as sharing Sorkin's goal. At one point, he says:

Bankman-Fried: I think I have a duty to talk to people. I have a duty to explain what happened. And I think I have a duty to do everything I can try to try and do what's right. If there is anything I can do to try and help customers out there. And I don't see what good is accomplished by me just sitting locked in a room pretending the outside world doesn't exist [like my lawyers advised me to do].

(Sorkin 2022)

About two weeks before this interview, Bankman-Fried gave an interview in which he claimed that in public, he merely pretends that he cares about ethics. Sorkin had read the interview. He even briefly mentions this interview to Bankman-Fried. Here's a representative excerpt:¹³

Piper: So the ethics stuff—mostly a front? People will like you if you win and hate you if you lose and that it's how it all really works?

Bankman-Fried: Yeah. I mean that's not all of it. But it's a lot....

Piper: You were really good at talking about ethics, for someone who saw it all as a game of winners and losers.

Bankman-Fried: Yeah. Hehe. I had to be. It's what reputations are made of, to some extent. I feel bad for those who get fucked by it.

(Piper 2022)

I do not think that any point in this interview, Sorkin believed that Bankman-Fried genuinely intended to cooperatively pool information, about his financial crimes, with Sorkin. It is likely that Bankman-Fried knew that Sorkin was skeptical of him. Nonetheless, Bankman-Fried clearly aimed to *appear* to cooperate with Sorkin's stated goals. The interview is funny, in part, because Bankman-Fried so dramatically fails to actually seem cooperative.

How should we interpret what Bankman-Fried does? To interpret him, we have to keep track of what Bankman-Fried is presupposing about himself and about the conversation. And though we don't believe his sincerity, Bankman-Fried is clearly and publicly presupposing that he is cooperating with Sorkin.

Here's what I am going to do now. As a warm-up, I will show how the Gricean ought to explain cases of straightforward, failed deceptive *misleading*. To do this, we will need to revisit Grice's theory of utterance interpretation. After that, I will show how the Gricean can generalize this discussion in order to explain cases in which both parties more or less jointly know that they are not cooperating. That is, she can generalize it to

¹³ Bizarrely, Bankman-Fried conducted this interview via Twitter DM. For readability, I have lightly edited both Piper and Bankman-Fried's text.

cases in which parties know and know that know that they know that they know...that they are not cooperating (Lewis 1969).

a. Failed deceptive misleading

By Grice's own lights, his analysis of conversational implicature in "Logic and Conversation" is incomplete. He is missing a crucial piece of the explanatory puzzle—albeit one that doesn't seem to matter, *until* we confront non-cooperative cases.

Rewind all the way back to **Bank Account 1**. In that vignette, Sam's lawyer, Rose, recognizes that Sam has conversationally implicated that he has never had a personal bank account in Hong Kong. To do this, Rose inferred (i):

1. Sam did not only mean what he said.

How did she do this? She reasoned like this:

2. Sam and I aim to cooperatively pool our information about Sam's finances.
3. Given that Sam knows the answer to my question and given (2), Sam is rationally required to answer my question.
4. If Sam just meant what he said, he wouldn't have answered my question.

Here's the step we (and Grice) were missing:

5. Sam **intends for me to infer** what he means through reasoning about the information in (2-4).

Consider Grice's theory of meaning (Grice 1957, 1989). For Grice, what we mean is wholly determined by a special set of our intentions—that is, our meaning intentions.¹⁴ To recognize what Sam means, Rose must recognize the content of Sam's meaning intentions. Moreover, at least one of Sam's intentions is a recognition intention: Sam intends for Rose to recognize *that he intends for her to recognize his meaning intentions*.

That alone doesn't require the step in (5). What requires the step in (5) is Grice's innovation in a later paper (Grice 1989 [1969]: 164). Grice observed that when a speaker means p , she doesn't just intend for her addressee to come to know that she means p . She intends for her addressee to follow a specific inferential path in order to achieve this knowledge. Call this the speaker's *inferential pathway intention*.

To intuitively motivate the constraint, Ray Buchanan (2013) presents the following case. Due to a medical procedure, I can't talk. You ask me what movie I would like to see this evening. I draw what I think is a cowboy hat and show it to you. I intend for you to recognize that I have drawn a cowboy hat and thus infer that I want to see the cowboy movie *Flat-top Mountain*.

¹⁴ Here I should make explicit what I have been presupposing. In "Logic and Conversation," is Grice offering us a metaphysical theory or an epistemic theory of conversational implicatures? More precisely: is Grice offering us a theory about *what makes it the case* that a speaker has conversationally implicated p or about *how* the addressee comes to know that the speaker has conversationally implicated p ? In order to make Grice's work coherent, it must be the latter. Grice has a single answer to the metaphysical question: if a speaker means anything at all (including a conversational implicature), she does so in virtue of the fact that she has a certain suite of intentions. For more on this distinction within Grice's work, see Neale (1992, 2004, 2005, 2016) and Harris (2016).

You do infer that I want to see the cowboy movie *Flat-top Mountain*. So that part of my communicative plan is a success. But you misidentified my intended drawing. You believe that I intended to draw a mountain and intended you to infer, on that basis, that I want to see *Flat-top Mountain*. So part of my communicative plan fails. The fact that my ultimate aim succeeds, though my inferential pathway intention fails, is just a matter of dumb luck.

Both Grice and many Griceans believe that in such cases, you have failed to fully *understand* my communicative act (Grice 1989 [1969]; Schiffer 1972, 2016; Bach and Harnish 1982: 85-87; Buchanan 2013). To fully understand a speech act requires more or less conforming to the speaker's inferential pathway intention.

Some philosophers endorse a view that is even stronger than this. For Grice and for Stephen Schiffer, my inferential pathway intention is a *constitutive* meaning intention (Grice 1989 [1969]; Schiffer 1972, 2016). If you take the right inferential pathway, but you don't recognize that I intend for you to take that pathway, you don't fully know what I meant. For our purposes, we don't need to settle on the stronger or weaker version of this view. Either way, an addressee who wants to understand the speaker will aim to identify and conform to the speaker's intended inferential pathway for her.

This means that Grice himself has the resources to explain what's happening in cases of failed misleading. Consider Yuna and Sam's interview. Suppose that Sam reasonably believes that he's in a position to deceive Yuna about his Swiss bank accounts. For example, suppose that Sam's financial crimes are well-hidden, and suppose that Yuna interviews Sam under a false and flattering pretext. ("The world is just so curious about how you've become so successful!") So Sam knows that he is not fully cooperating with Yuna's desire to pool information about his financial situation. He also believes that Yuna hasn't yet realized this.

When Yuna interprets Sam's reply in (2b), Yuna reasons about how Sam *intends for* her to interpret him. Sam intends for Yuna to reason from her belief in (2) and (3) to the conclusion that Sam means more than what he said:

2. Sam and I aim to cooperatively pool our information about Sam's finances.
3. Given that Sam knows the answer to my question and given (2), Sam is rationally required to answer my question.

Since Yuna knows what Sam believes about what she believes, Yuna recognizes Sam's inferential pathway intention. Though Yuna knows that (2) and (3) are false, this is no barrier to Yuna following the pathway and thus understanding what Sam means. She can simply *assume* (2) and (3) and see what follows. If she does this, Yuna will thereby infer that Sam meant (2c) that he never had a personal Swiss bank account.

I want to pause to emphasize an important point. Yuna does not just infer that *if* Sam were cooperative, then he *would* implicate (2c). She infers that Sam *did* implicate (2c). Yuna's inference is sound, even though she knows Sam is not, in fact, cooperative in the relevant sense. This is because Yuna ultimately aims to recognize Sam's meaning *intentions*. It's just that Sam intends for Yuna to recognize these intentions by reasoning under false assumptions.

Still, an analogous case from the theory of reference may help make the point more intuitive.¹⁵ Suppose that Elena is a member of the Board of Trustees of a major university. She knows that the Board has just voted to replace the university president, Lizzie Jones, with Bob Smith. The change is effective immediately, but the Board does not plan to make the vote public until the next day.

¹⁵ The example that I discuss below is a more complicated variation of Donnellan's famous man with the martini (Donnellan 1966: 55). My discussion is heavily indebted to the discussion in Stalnaker (2002: 717-718).

That night, Elena takes an interview with a journalist, Arthur. Believing that Arthur is ignorant of the secret vote, Elena says (3a). She intends to mean (3b):

University interview

Elena: 3a. I think that the current university president's view on free speech is a good one.
 3b. I think that Lizzie Jones' view on free speech is a good one.

In fact, Arthur is secretly good friends with a member of the Board of Trustees. That trustee has already told Arthur the news about Lizzie Jones, and so Arthur knows that Elena knows that the president is Bob Smith. Nonetheless, Arthur knows what Elena believes about Arthur's knowledge. So Arthur knows that Elena intends for him to reason as follows:

1. The current university president is Lizzie Jones.
2. Therefore, Elena is referring to Lizzie Jones.

Arthur infers that *if* Elena believed that they both believed that the university president were Lizzie Jones, *then* Elena would be referring to her. Arthur knows that Elena knows that the president is not Lizzie Jones. Nonetheless, he concludes that Elena *did* refer to Lizzie Jones, and Elena *did* mean (3b). Arthur arrived at knowledge about what Elena meant, even though he knowingly adopted a false assumption to get there.

In both cases, Arthur and Yuna are relying on the same shared assumption about the speakers:

Intelligibility

If a speaker means *p*, she intends for the addressee to know that she means *p*.

As we've seen, this constraint drops out of Grice's theory of meaning. If I believe that you mean *p*, then given the very nature of meaning, I believe that you intend for me to know that you mean *p*.

Nonetheless, there is one important difference between the two cases. Yuna, unlike Arthur, is *also* relying on her knowledge that she and Sam are *not* actually cooperatively exchanging information. Yuna knows that Sam wants to deceive her. It is, in part, this knowledge that enables Yuna to recognize that Sam aims to appear to cooperate with her. And this knowledge enables Yuna to recognize Sam's inferential pathway intention.

As Jessica Keiser points out about cases like these, "it is the assumption of conflict that allows the [addressee] to calculate [the speaker's] implicature" (Keiser 2022: 51). But that doesn't show that Grice's appeal to the Cooperative Principle was misguided. Rather, it shows that Sam and Yuna can *use* the Cooperative Principle as an interpretational resource, even when they both know that they do not aim to cooperatively pool their knowledge. They can do this, because they each know what Sam aims to do. He aims to appear to cooperate with Yuna. And so, what he does will bear a systematic relationship to what he would do, were he cooperative.

b. The generalized account

The sense in which Sam aims to *appear* to cooperate with Yuna is clear and intuitive. Sam's aims are deceptive. He wants Yuna to believe that he has nothing to hide, that he's an honest man and an open book, that she can trust what he says and does in the interview. So Sam intends to act in a way that gives Yuna (further) reason to believe that he aims to cooperatively share information with her. And Sam intends to avoid acting in a way that would give Yuna any (further) reason to believe that he does not have this aim.

We can say something similar about Bankman-Fried and Sorkin. Bankman-Fried is trying to deceive Sorkin and anyone who watches or reads the interview. He's just (hilariously) trying and failing.

But even if you don't intend to deceive your interlocutor, it's possible to aim to appear to cooperate with her. In fact, it's possible to aim to appear to cooperate with your interlocutor and intend *not* to deceive her about your actual aims.

Consider this story. Betty's philandering husband, Don, has been having an affair with another woman, Faye.¹⁶ Don is barely trying to hide the evidence. Moreover, because he wants Betty to know how much disdain he has for her, he wants Betty to know that he's barely trying to hide the evidence from her.

After Don has been out all night, Betty asks him, in a faux-casual voice:

Blatant affair

Betty:	4a. Have you seen Faye lately?
Don:	4b. She's had Covid.

Don is conversationally implicating:

4c. I haven't seen Faye lately.

Don intends for Betty to recognize that his implicature in (4c) is insincere, and he intends for her to recognize this very intention. That is the mechanism through which he is expressing his disdain for her. On an intuitive level, we understand what Don is doing, and we know that he's doing something morally perverse.

But when explicitly articulated, Don's intention can seem exotic. Don intends for Betty to recognize that he means (4c). So Don intends for Betty to recognize that he's aiming to appear to cooperatively exchange information with her. But Don also intends for Betty to recognize that he is implicating (4c) insincerely. So Don intends for Betty to simultaneously recognize that he is not in fact cooperatively pooling his belief with hers and that he intends for her to recognize this fact. What is the relevant notion of "appearance" here?

It will help to have some theoretical terminology. To describe my account, I will use the popular common ground framework that Robert Stalnaker has developed (Stalnaker 2002, 2014). Because this framework presupposes a *psychologistic* theory of conversational dynamics—that is, it explains conversations primarily in terms of the intentions and mental states of its participants—it meshes seamlessly with Grice's view. (Stalnaker himself said that Grice had a "profound influence" on his work (Stalnaker 2014: 1).)

The basics of the view are well-known, so I will review them quickly.

Stalnaker persuasively argued that whenever we converse with each other, we take a chunk of information for granted. This information is *intersubjective* and *shared*. It serves as a kind of mutually accepted background for the conversation. As the conversation proceeds, this body of information evolves. In the simplest case, we take turns adding information to it, so that it becomes richer and richer.

Stalnaker calls this body of information the *common ground*. When all goes well in a conversation, the common grounds guides and explains what we do and how we interpret each other. For the theorist, the common ground also has a *rationalizing* role: it explains why interlocutors are behaving in a reasonable or at least intelligible way. Because of this, the common ground can help us illuminate the relationship between what a speaker says, what the speaker means, and what inferential pathway the speaker intends for the addressee to follow to identify both acts.

¹⁶ This is a mash-up of two different cases. I've modeled the dialogue on the example in Asher (2012: 48), for which he credits Potts and Stone (p.c.). I've modeled the scenario after the one in Meibauer (2014), who in turn modeled Don and Betty after characters in the HBO show *Mad Men*.

In the simplest conversation, what is common ground is what we jointly know. When we talk, we take turns asserting what we know. Since we jointly know that we trust each other, our pool of joint knowledge keeps expanding.

But obviously, most conversations are not like these ones. In order to effectively model real-world conversations—which are full of error, deceit, pretense, assumption, and “going along to get along”—Stalnaker identifies the common ground with what we jointly *accept* for the sake of the conversation:

Common ground

Some proposition p is **common ground** in a conversation c iff all participants in c :

- (i) accept (for the sake of c) that p ,
- (ii) accept (for the sake of c) that (i),
- (iii) accept (for the sake of c) that (ii),
et cetera...

(Stalnaker 2014: 4, 25)

To accept p for the sake of the conversation is to treat p as true for the sake of the conversation—or something more or less like that (Stalnaker 1984: 84, Stalnaker 2002: 716). There are different ways to accept p . One way is to believe or know p . Another is to assume or hypothesize or suppose p (Stalnaker 2002: 717, 2014: 25). A third way is to pretend p (Stalnaker 2002: 718).

Thus, sometimes when we talk, it is common ground that p , even though one or both of us fails to believe p . Maybe you insist, “I love your stew!” but you’re lying to me. In that case, you’ve accepted that you love the stew in virtue of pretending that you do, but I think you’ve accepted it in virtue of believing it. Or maybe we’re assuming that the university will radically slash our budget next year, in order to plan for the relevant contingencies. In that case, we jointly know that we’ve merely assumed this proposition.

Other times, when we talk, we jointly believe or know p , but p is not common ground (Stalnaker 2014: 46-47). In public, we might insist that we’re happy to see each other, that we recently liked each other’s conference presentation, and that we wish the other well—even while we jointly know that we are locked in a vicious feud and that none of what we are saying is true. Or suppose that everyone in a company jointly knows that the boss is illegally embezzling corporate money. To avoid retaliation and protect their own interests, employees might express feigned puzzlement about “where the money is going.”

To *presuppose* p is to take it for granted that all parties are taking p for granted in the conversation:

Presupposition

A participant in a conversation c presupposes p for the purposes of c iff she accepts that p is common ground in c .

(Stalnaker 2014: 4, 25)

For our purposes, we now just need to make simple observation. Propositions about what we aim to do in a conversation are propositions like any other. This means that we can accept propositions about what we aim to do in a conversation, even if we do not believe or know these propositions. And it means that we can accept propositions about we aim to do in a conversation, even when these propositions are incompatible with what we individually or jointly know to be the case.

For example, we can presuppose that we primarily aim to make light and breezy small talk with each other—even when we jointly know that we each *actually* primarily aim to show off to our advisor, who is sitting next to us. We can presuppose that you are telling me about your vacation because I genuinely want to hear about it—even if you are unsure whether I actually care or whether I am just asking out of habitual politeness. We can presuppose that we each aim to sincerely tell each other what we think about each other's work—even if we jointly know that we are just making polite chit-chat, and neither of us intends to say what she *really* thinks.¹⁷

In such cases, the participants have what I will call *apparent goals*:

Apparent goal (presupposition)

If a speaker presupposes that she has some goal g in a conversation c , then g is one of the speaker's apparent goals.

Apparent goal (common ground)

If the common ground of some conversation of c entails that its participants have some joint goal g , then g is an **apparent goal** of c .

A speaker's apparent goal is *merely* apparent when she does not actually have it. A conversational goal is *merely* apparent when the conversational participants do not really share it. But not all apparent goals are merely apparent. Suppose that you and I are having a simple and straightforward conversation, in which we are talking about shark attacks. Since we are appropriately coordinating, it is common ground that we jointly aim to pool our knowledge about shark attacks. So that goal is apparent. But it is only an apparent goal because we *actually do* share the goal.

With all this in mind, here is how I propose that we reform Grice's account. First, whenever a speaker makes a conversational implicature, she *appears* to cooperate with her interlocutor. Or, to put it more precisely:

Apparent Cooperation Thesis

Whenever a speaker makes a conversational implicature, she is presupposing that she aims to cooperatively exchange information with her interlocutor.

To be clear, I am not suggesting that *every time* we talk to each other, one or both of us appear to cooperatively exchange information with each other. Here's a perfectly intelligible conversation:

Trash talk

Kaz:	Screw you.
Liam:	No, screw <i>you</i> .
Kaz:	Seriously, fuck yourself.
Liam:	You go fuck yourself, you piece of shit!

There's not even an appearance of cooperative information exchange here—not unless you think that exchanging insults constitutes a cooperative shared project. But I also don't think that Liam and Kaz *can* make conversational implicatures to each other.

¹⁷ If we pretend that we share a conversational aim that we know that we do not share, it does not follow that everything we do within that conversation is a pretense. Similarly, if Arthur and Elena jointly pretend that Lizzie Smith is the current president of the university, it does not follow that they are merely *pretending* to have an interview and *pretending* to make assertions within it. This means that the conversations that we are discussing lack a homogenous “conversational tone” (Yalcin 2007: 1008). Some propositions are common ground in virtue of the fact that we jointly believe them; others are common ground because we jointly pretend that they are true; and so on.

Second, within this framework, we can derive what I will call the Apparent Cooperative Principle:

Apparent Cooperative Principle

Act so that **the common ground** entails that you have made your contribution such as is required, at the stage at which it occurs, by the **apparent goal** of the talk exchange in which you are engaged.

In other words, the Apparent Cooperative Principle requires us to act so that the common ground entails that we have conformed to the Cooperative Principle.

As we have seen, conforming to the Cooperative Principle generally requires conforming to the Four Maxims of Quality, Quantity, Relevance, and Manner. So if we aim to conform to the Apparent Cooperative Principle, we are generally required to act so that the common ground entails that we are conforming to Quality, Quantity, Relevance, and Manner. (Note that this is true, whether or not these Maxims are ultimately social norms.)

To understand how the view works, let's first consider simple and cooperative conversations. In such cases, we will generally have reason to conform to the Apparent Cooperative Principle insofar as we have reason to conform to the Cooperative Principle. In these conversations, our apparent goal is just our actual goal, and what is common ground is just what we jointly know or believe.

So the Apparent Cooperative Principle will require this:

Apparent Cooperative Principle (Simple case)

Act so that you jointly know that you have made your contribution such as is required, at the stage at which it occurs, by the actual goal of the talk exchange in which you are engaged.

This is, strictly speaking, stronger than what Grice's Cooperative Principle requires. Grice's original Principle only requires us to *actually* cooperatively pursue our conversational aim. My Principle requires us to jointly *know* that we are cooperatively pursuing our aim.

That is because I am assuming that when we are cooperatively conversing towards some goal, we are required to *show* that we are being cooperative. We must act in ways that make our cooperation mutually manifest to all participants.¹⁸ If you mention some fact about whale attacks, but I don't see how it relates to our discussion about shark attacks, then something has gone wrong—even if what you said *was* perfectly relevant, and I was simply too distracted or confused to realize how.

Now let's consider our non-cooperative cases, in which one or more speakers *merely* aims to appear to cooperate with her interlocutors. Let's split these cases into two groups.

First, suppose that like the fictional Sam and the actual Sam Bankman-Fried, you aim to deceive your interlocutor about the extent to which you are cooperating with her. Since you aim to deceive your interlocutor, you must intend for her to falsely *believe* that she is in a simple and cooperative conversation. You must intend for her to falsely believe that you jointly know and therefore jointly presuppose that you are cooperatively exchanging information. And for that to work, you must act in ways that ensure that the common ground entails that you are cooperating with her. So insofar as you aim to deceive your interlocutor in this way, you are rationally required to conform to the Apparent Cooperative Principle.

Second, suppose that like Don, you do not aim to deceive your interlocutor about the extent to which you are cooperating with her. Nonetheless, you aim to construct and act within a gossamer-thin conversational

¹⁸ Interpreting contemporary Gricean pragmatics, Szabó explicitly endorses a similar constraint: “Ensure it is common knowledge how speech acts in conversation contribute towards the goal of the conversation” (Szabó 2020: 61).

pretense. Through engaging in this pretense, you may, like Don, intend to express some sentiment or commitment. Or you may simply want to conform to norms of good politeness and decorum. (We will consider more of these cases in the next section.) To construct and engage in this kind of pretense just is to presuppose the pretense throughout the conversation. So insofar as you aim to engage in this kind of pretense, you are rationally required to conform to the Apparent Cooperative Principle.

So we have three cases: the simple and cooperative case, the deceptive pretense case, and the non-deceptive pretense case. In all three of these cases, the speaker intends for her addressee to *know* that she aims to appear to cooperate with her interlocutor. After all, in all three of these cases, the speaker wants it to become common ground that the speaker aims to appear to cooperate with her interlocutor. And to do that, the addressee must generally know what the speaker is presupposing and presuppose it herself.

What this means, then, is that in all *three* cases, the speaker intends for her addressee to know that the speaker is rationally required to conform to the Apparent Cooperative Principle. Thus, we arrive at this new version of Grice's Key Assumption:

Key Assumption 3

Whenever a speaker successfully makes a conversational implicature, the addressee reasonably believes that if the speaker had not made that conversational implicature, she would have violated the Apparent Cooperative Principle.

And thus, we have arrived at a fully unified analysis of conversational implicature.

IV. Other examples of merely apparent cooperation

I think that the best way to advertise the strengths of my account is to show how it handles different kinds of cases. In this final section, I will do just that.

a. "Protesting my innocence" cases

In one class of cases, the speaker insists that she is cooperating with her addressee, even while the addressee refuses to accept that she is doing so. For example, consider this conversation between Sam and Yuna:

Bank account 3

Yuna:	So you're saying that you never had a Swiss bank account?
Sam:	Yes, exactly.
Yuna:	But that's not what these records show. I'm afraid it's just obvious that you've been lying to me this whole time.
Sam:	No, no, you've got it all wrong. See, look at these documents here...
Yuna:	5a. Now you're trying to distract me with this paperwork. Just admit that you once had a Swiss bank account.
Sam:	5b. My company did!

In saying (5b), Sam is clearly implicating:

5c. And I did not have a personal Swiss bank account.

Isn't this a case in which Sam and Yuna lack even the *appearance* of cooperating with each other? And so doesn't this show that my analysis doesn't fully solve the problem?

In one respect, we can answer “yes” to the first question. In **Bank Account 3**, it is not common ground that Sam aims to cooperatively exchange information with Yuna. Yuna accuses Sam of lying and trying to distract her. So Yuna is refusing to presuppose that Sam aims to cooperatively exchange information with her. But this isn’t a problem for my account. If anything, my account predicts it. Sam and Yuna are just in a bog-standard conversational crisis.

We often squabble about what we want to become common ground, in exactly this sort of way:

Royal sighting

Karen:	I just saw the King of Iceland.
Alexei:	There is no King of Iceland. That was just Gunnar.
Karen:	Yeah, I was super surprised at how young the King is, actually.
Alexei:	What are you talking about?
Karen:	How do you not know about the King of Iceland?

When she first speaks, Karen presupposes that there is a King of Iceland, and Alexei refuses to accept this presupposition. But even though Karen knows that Alexei refuses to accept her presupposition, Karen continues to presuppose it. It’s not common ground that Gunnar is the King of Iceland. But it is common ground that *Karen is presupposing* that Gunnar is the King of Iceland. This is why Alexei can interpret her so easily.

Similarly, in **Bank account 3**, it’s common ground that *Sam is presupposing* that he and Yuna actually are cooperatively pursuing a shared project. In order to identify Sam’s implicature, Yuna recognizes that Sam intends for her to temporarily assume what he is presupposing about himself. Because Sam’s presuppositions are common ground, Yuna has no problem doing this.

In both cases, the conversation has come to a screeching halt, because the participants cannot agree on how to proceed. They cannot reach agreement, but they will not agree to disagree either. Nonetheless, because it is common ground *what* the conversational crisis is about and which side each party has taken within it, they remain intelligible to each other.

b. “For the sake of audiences” cases

In some cases, the speaker aims to deceive her audience, not her addressee. But in order to do successfully do that, the speaker aims to appear to cooperate with her addressee. This is a frequent occurrence in courtrooms: the guilty defendant is speaking to the prosecutor, but she aims to deceive the jury. It is often a tactic within political debates and negotiations: neither candidate aims to deceive the other, but each aims to deceive her constituents or the moderator or the press. And it almost certainly happens all the time online.

Here’s how one anonymous online commenter put it:

One of the valuable parts of debating [online] is that you are aiming to convince the audience - the silent lurkers who are watching. In that situation, you are a representative of your position, and your opponent of his. So if he comes screaming at you, insulting you, slamming walls of text, and you respond politely, reasonably, even kindly, it is going to be a much better way to advertise to your audience that your position is more reasonable.

(Aetole 2022)

Suppose that Henry and Charlotte are feuding on social media about Trump's latest policy. When she replies to Henry, Charlotte keeps insisting that she wants to have a "good faith" discussion with Henry, that she's only asking reasonable questions, and that the statistics that she is posting are utterly germane to their discussion. In fact, Charlotte is steering the conversation towards topics that she knows will provoke Henry into an explosive reaction. And she's posting phony and irrelevant statistics. Charlotte doesn't care whether Henry recognizes Charlotte's true goals; she only cares that silent onlookers believe that she is cooperating.

In these cases, the addressee often manages to identify the speaker's presuppositions *because* this audience is so salient. Consider the original "Swiss bank accounts" case, which many philosophers have discussed at length.¹⁹ Suppose that Maya is a prosecutor, and Sam is a defendant on the stand. They have the following exchange:

Bank account 4

Maya:	6a. Have you ever had a bank account in Switzerland?
Sam:	6b. My company had one there for about six months.

Some philosophers have insisted that even on the stand, Sam conversationally implicates (6c):

6c. No, I never had a personal bank account in Switzerland.

Sam wants the jury to believe that he is an innocent man, that he has nothing to hide, and he is answering Maya's questions in a spontaneous and helpful way. And the best way to do that, in this context, is to presuppose, within his conversation with Maya, that he is cooperating with her. Maya can identify Sam's presupposition and, thus, Sam's implicature in (6c) *because* it's so mutually obvious that a defendant like Sam would aim to deceive the jury in this way.

c. Ritual conversations

Finally, in a special class of cases, speakers and addressees actually work together to maintain the mere appearance that they aim to cooperatively exchange information with each other. In successfully working together in this way, the participants either conform to background social norms, or indirectly express some attitude towards each other,²⁰ or both. I call these kinds of conversations *ritual conversations*.

Consider this completely familiar interaction between two professional philosophers at a conference:

Conference talk

Kaz:	7a. Lucy, it's so great to run into you. I admired your latest article. Always so eye-opening.
Lucy:	7b. I'm flattered you read it. But I'm really interested to know what you've been thinking about recently. You should tell me sometime.

What is happening here depends upon minutiae: the precise contours of Kaz and Alexei's relationship, the exact shared knowledge they have about each other, and the subtle intonations and body language that they are using.

¹⁹ See footnote 7.

²⁰ Interpreting the work of the sociologist Erving Goffman (1959, 1967, 1971), Berstler (forth. a) calls this *dramatic communication*. The following discussion is indebted both to Goffman and to the discussion in Berstler. For a technical discussion of how we can coordinate on information even when that information is not common ground, see Berstler (forth. b).

But it's completely believable (and, indeed, a familiar scenario) that Kaz does not intend for Lucy to believe that he *really* admires her paper. And he intends for Lucy to recognize this intention. Similarly, Lucy does not intend for Kaz to believe that she is all *that* interested in his latest philosophical thoughts. And she intends for Kaz to recognize this intention. If that's right, what are they doing?

Maybe they each want the other to infer nothing about their actual opinions. In saying (7a) and (7b), Kaz and Lucy only want to express that they *respect* and *esteem* each other as colleagues. Because of this, Kaz should not expostulate at length about his recent views; for all he knows, Lucy doesn't want to hear him about them. And Lucy should not ask Kaz what he found so "eye-opening" about her article; for all she knows, when pressed, Kaz will have nothing substantive to say about it. And though she would be expressing a true and sincere sentiment, it *definitely* wouldn't be cooperative for Lucy or Kaz to openly say:

7c. Aww, I don't know what your views really are, because I know that we're doing a little social ritual thing, but I appreciate that you respect me!

Or maybe Lucy and Kaz each want the other to know that they *don't* believe what they say in (7a) and (7b), but they are willing to say it anyway. Maybe they jointly know that they each loathe each other and hate each other's work. But they are saying (7a) and (7b) in order to conform to professional standards, or in order to indirectly communicate to each other that they will not openly harm the other's professional interests, or both. Once again, though she would be expressing a true and sincere sentiment, it wouldn't be cooperative for Lucy or Kaz to openly say:

7d. I know you're just pretending, but I'm glad that we can talk like this, and that even though you hate me, you're not going to try to black ball me or anything.

In fact, if Lucy or Kaz *did* say either (7c) or (7d), that would undermine the communicative success of their social ritual.

So in both cases, Lucy and Kaz are cooperating insofar as they jointly aim to *merely appear* to aim to pool their knowledge about what they each think of each other's work. These cases show that merely apparent goals are not necessarily the calling card of non-cooperative conversations. We sometimes have merely apparent goals, even when and because we are pursuing a substantial shared project within the interaction.

V. Conclusion

Buried deep within Grice's "Retrospective Epilogue," Grice admits that his view faces a Non-Cooperation Problem. His remarks on the problem are brief, perfunctory, and a little bit mysterious:

Moreover we have to remember to take into account a secondary range of cases like cross-examination in which even the common objectives are spurious, apparent rather than real; the joint enterprise is a simulation, rather than an instance, of even the most minimal conversational cooperation; but such exchanges honor the cooperative principle at least to the extent of aping its application. A similarly degenerate derivative of the primary talk-exchange may be seen in the concerns spuriously exhibited in the really aimless over-the-garden-wall chatter in which most of us from time to time engage.

(Grice 1989: 369-370)

There is a way in which I have tried to vindicate this passage. I have argued that in the problem cases, the speaker aims to appear to cooperate with her interlocutor. Because of this, the speaker is rationally required

to *appear* to conform to the Cooperative Principle. Moreover, the addressee is in a position to recognize these facts and so reason about what the speaker *would* do, were she cooperative. In that sense, the speaker and addressee “honor” the cooperative principle in “aping its application.”

But I have also re-imagined the Gricean framework so that it presents a unified framework for theorizing about conversational implicatures. The trick was that I wheeled in talk about the common ground. And I argued that when we actually aim to cooperatively exchange information *and* when we merely aim to appear to do so, we must act so that the common ground entails that we are conforming to the Cooperative Principle.

So, “non-cooperative” or “non-ideal” conversations don’t pose some radically intractable problem for the *spirit* of the Gricean project. There may be deep other problems with the spirit of Griceanism, but the Non-Cooperation Problem isn’t one of them.

But here is where I do agree with many non-ideal philosophers of language. There is one sense in which the conversations that I have been discussing are “secondary to” or “derivative of” their cooperative counterparts. What it is to appear to aim to exchange information depends upon what it is to actually aim to exchange information. So in some sense, the structure of some non-cooperative conversations will depend upon the structure of their cooperative counterparts.

But that does not mean that non-cooperative conversations should be of secondary *theoretical* importance. Once we have a more precise way to theorize *about* them, we can more precisely theorize about the many sneaky, strategic, provocative, prosocial, and undoubtedly theoretically *interesting* things that speakers do within them.

Bibliography

Aetole. Comment on “About debating the Alt-Right, MRA’s, etc.” *Reddit.com*. 2022. Accessed 28 February 2025.

https://www.reddit.com/r/MensLib/comments/pu6a4g/about_debating_the_alt_right_mrass/etc/.

Anderson, Luvell and Michael Barnes. “Hate Speech.” *Stanford Encyclopedia of Philosophy*. 2022.
<https://plato.stanford.edu/entries/hate-speech/>.

Asher, Nicholas. “The Non-Cooperative Basis of Implicature.” Published proceedings of *Logical Aspects of Computational Linguistics* 7. Edited by Denis Béchet and Alexander Dikovsky. 45-57. New York: Springer, 2012.

Asher, Nicholas and Alex Lascarides. “Strategic Conversation.” *Semantics & Pragmatics* 6.2 (2013): 1-62.

Bach, Kent and Robert M. Harnish. *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press, 1979.

Berstler, Sam. “What We Won’t Say.” *EurAmerica: Social Meaning and Reality*. Edited by Hsiang-Yun and Quill R. Kukla. Forthcoming a.

-----, “The Structure of Open Secrets.” *Philosophical Review*. Forthcoming b.

Beaver, David and Jason Stanley. *The Politics of Language*. Princeton, NJ: Princeton University Press, 2023.

Bird, Graham. "Speech Acts and Conversation—II." *Philosophical Quarterly* 29.115 (1979): 142-152.

Bratman, Michael. *Shared Agency: a Planning Theory of Acting Together*. New York: Oxford University Press, 2014.

Brown, Penelope and Stephen C. Levinson. *Politeness: Some Universals in Language Use*. New York: Cambridge University Press, 1987 [1978].

Buchanan, Ray. "Reference, Understanding, and Communication." *Australasian Journal of Philosophy* 92 (2013): 55-70.

Camp, Elisabeth. "Insinuation, Common Ground, and the Conversational Record." In *New Work on Speech Acts*. Edited by Daniel Fogal, Daniel W. Harris, and Matt Moss. 44-66. New York: Oxford University Press, 2018.

Capellen, Herman and Joshua Dever. "On the Uselessness of the Distinction between Ideal and Non-Ideal Theory (at least in the Philosophy of Language)." In *Routledge Companion to Social and Political Philosophy of Language*. Edited by Justin Khoo and Rachel Stern. 91-105. New York: Routledge, 2021.

Davis, Wayne. *Implicature: Intention, Convention, and Principle in the Failure of Gricean Theory*. New York: Cambridge University Press, 1998.

Donnellan, Keith. "Reference and Definite Descriptions." *Philosophical Review* 75 (1966): 281-304.

Ellefson, Gretchen. "Conversational Cooperation Revisited." *Southern Journal of Philosophy* 59.4 (2021): 545-571.

Englehardt, Jeff and Molly Moran. "Using the Ideal/Nonideal Distinction in Philosophy of Language (and Elsewhere)." *Australasian Journal of Philosophy*. Forthcoming.

Goffman, Erving. *Relations in Public*. New York: Basic Books, 1971.

-----. *Interaction Ritual: Essays on Face-to-Face Behavior*. New York: Random House, 1967.

-----. *The Presentation of the Self in Everyday Life*. New York: Anchor Books, [1956] 1959

Green, Mitchell S. "Quantity, Volubility, and Some Varieties of Discourse." *Linguistics and Philosophy* 18 (1995): 83-112.

Grice, H.P. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press, 1989.

-----. "Meaning." *Philosophical Review* 66 (1957): 377-388.

Harnish, Robert M. "Logical Form and Implicature." In *Pragmatics: a Reader*. Reprint. Edited by Stephen Davies. 316-364. New York: Oxford University Press, 1991 [1976].

Harris, Daniel W. "Intentionalism and Bald-Faced Lies." *Inquiry*. 2020.

-----. "Intentionalism versus the New Conventionalism." *Croatian Journal of Philosophy* 16 (2016): 173-201.

Hesni, Samia. "Intuitions and Deference in Non-Ideal Philosophy of Language." In *The Routledge Handbook of Non-Ideal Theory*. Edited by Hilkje Hänel and Johanna Müller. 415-427. New York: Routledge, 2024.

Holdcraft, David. "Speech Acts and Conversation—I." *Philosophical Quarterly* 29.115 (1979): 125-141.

Hornsby, Jennifer. "Disempowered Speech." *Philosophical Topics* 23 (1995): 127-147.

Meibauer, Jörg. "Bald-Faced Lies as Acts of Verbal Aggression." *Journal of Language Aggression and Conflict* 2 (2014): 127-150.

Kasher, Asa. "Gricean Inference Reconsidered." *Philosophica* 29 (1982): 25-44.

Keenan, Elinor Ochs. "The Universality of Conversational Postulates." *Language in Society* 5.1 (1976): 67-80.

Keiser, Jessica. *Non-Ideal Foundations of Language*. New York: Routledge, 2022.

Kukla, Quill R. "Uptake and Refusal." *Inquiry*. Forthcoming.

----- "Performative Force, Convention, and Discursive Injustice." *Hypatia* 29 (2014): 440-457.

Langton, Rae. "Speech acts and Unspeakable Acts." *Philosophy and Public Affairs* 22 (1993): 293-330.

Lewis, David. *Convention: A Philosophical Study*. Cambridge, MA: Blackwell, 1969.

Maitra, Ishani. "Silencing Speech." *Canadian Journal of Philosophy* 39 (2009): 309-338.

Michaelson, Eliot. "What is the Proper Function of Language?" *Inquiry* (2024): 1-24.

Mühlebach, Deborah. "Non-Ideal Philosophy of Language." *Inquiry* 10 (2022); 4018-4040.

Mühlebach, Deborah and Nikki Ernst. "Discursive Resistance in a Non-Ideal World." In *The Routledge Handbook of Non-Ideal Theory*. Edited by Hilkje Hänel and Johanna Müller. 126-137. New York: Routledge, 2024.

Neale, Stephen. "Silent Reference." In *Meaning and Other Things: Essays in Honor of Stephen Schiffer*. 229-342. New York: Oxford University Press, 2016.

----- "Pragmatism and Binding." In *Semantics versus Pragmatics*. Edited by Zoltán Gendler Szabó. 165-285. New York: Oxford University Press, 2005.

----- "This, That, and the Other." In *Descriptions and Beyond*. Edited by Anne Bezuidenhout and Marga Reimer. 68-182. New York: Oxford University Press, 2004.

----- "Paul Grice and the Philosophy of Language." *Linguistics and Philosophy* 15.5 (1992): 509-559.

Nowak, Ethan. "Sociolinguistic Variation, Slurs, and Speech Acts." *Journal of Philosophy*. Forthcoming.

Piper, Kelsey. "Sam Bankman-Fried Tries to Explain Himself." *Vox*. 16 November 2022. Accessed 27 February 2025. <https://www.vox.com/future-perfect/23462333/sam-bankman-fried-ftx-cryptocurrency-effective-altruism-crypto-bahamas-philanthropy>.

Schiffer, Stephen. "Grice Semantics and Reference: Responses to Anita Avramides, Stephen Neale, and Kent Bach." In *Meaning and Other Things: Essays on Stephen Schiffer*. Edited by Gary Ostertag. 493-528. New York: Oxford University Press, 2016.

-----. *Meaning*. New York: Oxford University Press, 1972.

Solan, Lawrence, and Peter Tiersma. *Speaking of Crime: The Language of Criminal Justice*. Chicago: University of Chicago Press, 2005.

Sorkin, Andrew Ross. "Transcript of Sam Bankman-Fried's Interview at the DealBook Summit." *New York Times*. 1 December 2022. Accessed 27 February 2025. <https://www.nytimes.com/2022/12/01/business/dealbook/sam-bankman-fried-dealbook-interview-transcript.html>.

Stalnaker, Robert. *Context*. New York: Oxford University Press, 2014.

-----. "Common Ground." *Linguistics and Philosophy* 25 (2002): 701-721.

-----. *Inquiry*. Cambridge, MA: MIT Press, 1984.

Stokke, Andreas. "Truthfulness and Gricean Cooperation." *Grazer Philosophische Studien* 93 (2016): 489-510.

Stotts, Megan Henricks. "Conversational Maxims as Social Norms." *Inquiry* 67 (2024): 3087-3109.

Szabó, Zoltán Gendler. "The Goal of Conversation." *Aristotelian Society Supplementary Volume* 94.1 (2020): 57-86.

Thomason, Richmond H. "Accommodation, Meaning, and Implicature: Interdisciplinary Foundations for Pragmatics." In *Intentions in Communication*. Edited by Philip R. Cohen, J. Morgan, and M. Pollack. 325-364. Cambridge, MA: MIT Press, 1990.

Tirrell, Lynne. "Genocidal Language Games." In *Speech and Harm: Controversies over Free Speech*. Edited by Ishani Maitra and Mary Kate McGowan. 174-221. New York: Oxford University Press, 2012.

Yalcin, Seth. "Epistemic Modals." *Mind* 116 (2007): 983-1026