Chapter Three:

Planning Conversations Together

Daniel W. Harris*

Communicative acts typically aren't performed in isolation. They occur within larger communicative exchanges, in which participants take turns building toward shared goals. Conversations are sometimes quite organized, and there are benefits to this organization. We can set things up early in a conversation for later payoffs by introducing information or questions that shape downstream conversation in a way that allows for more impressive communicative achievements than we could otherwise manage.

Pragmatics has delivered us a number of successful models of how some of these phenomena work. Here, I will discuss three case studies: Roberts' (2012b) theory that conversations are organized around "questions under discussion" (§4), Grice's (1975) theory of conversational implicature (§5), and a number of related models of how contextual parameters that are set early in a conversation are permitted to evolve later on (Barker, 2002; MacFarlane, 2016, 2020) (§5). My overarching claim in this chapter is that we can understand the subject matter of these models, and synthesize their insights, by interpreting all of them as giving us partial glimpses at the shared "conversation plans" that we use to build and use to organize our conversations.[1] My theory of conversation plans will be built on top

*This is a draft of the third chapter of the book I am working on, whose working title is *Human Communication*. Drafts of other chapters can be found at `https://danielwharris.com/book/`. For feedback on earlier versions, I thank Ray Buchanan, Juan Vargas, Elmar Unnsteinsson, and the students in my Fall 2023 Graduate Seminar at the CUNY Graduate Center. I have not yet been able to act on all of their very helpful advice, but I am working on it.

[1] For some influential earlier discussions of conversation as a jointly planned activity, see Bratman (2014); Carlson (1982); Clark (1996); Cohen and Levesque (1985, 1990); Ginzburg (1995a,b,

of Bratman's (1992; 2014) theory of shared planning (§2). If I am right, then a capacity to build and work with shared plans is an important part of what it takes to be a pragmatically competent language user (§7). Before I get to these claims, however, I will introduce what I take to be another benefit of thinking of conversation in terms of shared plans—namely, that it gives us an important part of the explanation for why we form communicative intentions at all.

1 Why Reveal Our Intentions?

In Chapter 2, I gave an argument for thinking that a human communicator is typically guided by an intention to produce a certain state of mind in their addressee, which I called an "effective intention," and which is the first component of a communicative intention. In brief, my argument was that we must posit effective intentions as components of the hierarchical plans that link speakers' abstract aims to the specific bodily movements by which they seek to realize them. Effective intentions result from the planning process by which a communicator pairs an intended message with an intended addressee ("message design") and feature as important inputs to the process of reasoning about how to convey that message to that addressee ("signal design"). Together, these two halves of the communication-design process enable us to customize what we communicate and how we communicate it for our addressees, and this unlocks many communicatively valuable features of human communication and natural language.

In addition to an effective intention, a communicative intention also has a second component—the intention to reveal an effective intention to one's addressee, which I have called a "revelatory intention." My argument for positing effective intentions leaves open the question of why human communication also often involves revelatory intentions. Once I have an intention to change your state of mind in some way, why would I pursue this aim by trying to reveal this intention to you? What is the communicative point of making our intentions transparent?

Transparency is not *always* the best strategy. When my daughter is skeptical about a new food, I sometimes try to stoke a desire to try it by pretending that I don't want her to. I thereby conceal my effective intention in the service of achieving it. I

2012); Grice (1975); Grosz and Sidner (1986); Jankovic (2014); Perrault (1990); Roberts (2012b); Thomason (1990). I will attempt to absorb the lessons of several of these discussions below.

often try to get my students to have certain beliefs about a philosopher's ideas, not by telling them what to think, but by getting them to read the philosopher's work for themselves. In this case, I don't want the students to be thinking about what *I* intend them to believe about the text. Rather, I intend for them to arrive at the right beliefs by piecing together the textual evidence for themselves.

In many situations, these strategies would be very difficult to pull off. We rarely attempt reverse psychology, both because it is manipulative and because it works only in those unusual situations when we are faced with a contrasuggestible addressee. Likewise, we often lack the means to get someone to change their mind by confronting them directly with reasons for the change. In general, it is often difficult or impossible to change what someone thinks in a certain way without revealing the intention to do so. How would you get your barista to form an intention to make you an oat-milk latte without letting on that this is what you wanted? How would you get a friend to meet you at a certain place and time for lunch without letting them know that you intended them to do this? These are ordinary communicative tasks when we are being transparent about our intentions. If we sought to accomplish them while hiding our intentions, they would entail the sort of plotting and complex misdirection normally reserved for spy novels and surprise parties.

This line of thought makes it plausible that transparency about our intentions often accompanies easy and efficient communication. But why, exactly? Why, in particular, would revealing effective intentions so often be such a useful means to their satisfaction?

My answer, in brief, is that human communication goes much more smoothly when it is a rationally cooperative activity; activities of this kind work best when participants have a coherent shared plan about how to accomplish them; and building a coherent shared plan is almost impossible unless the planners make their intended contributions to this plan known to each other.

To see the basic point, it will help to compare conversation to a simpler example of joint action. Suppose that I would like to move my sofa to a very specific place on the other side of the room. The sofa is quite heavy, and so I will need your help. Now compare two variations on this scenario. In the first, I let you know where I intend to put the sofa, and I also try to make clear how I intend to contribute to achieving this goal (e.g. lifting from one side), and how I hope you will contribute (e.g. lifting from the other side). You are cooperatively disposed, and

so you form intentions that complement mine, thereby allowing us to efficiently coordinate our effort. In the second variation, I (for some reason) don't want you to know where I am trying to put the sofa, or what, exactly, I intend to do to accomplish my sofa-moving goal. In this version of the scenario, it *might* be possible to recruit your help in accomplishing my goal without your knowing what that goal is, but this bizarre form of success would presumably be considerably more difficult. In trying to imagine how it would go, I am struck by the image of two people yanking a sofa in different directions, or of one person expending considerable time and thought to cook up an elaborate scheme to manipulate the other's actions without revealing their true motives. Without knowing my goal, you could not be expected to efficiently invest your effort in pursuit of that goal. I would therefore have to compensate for your ignorance by investing far more effort into manipulating you. Compare these contrasting cases with the contrast between meeting a friend for lunch in the usual way and trying to manipulate them into arriving at the right time and place to their own surprise party.

Being a cooperative addressee, like being a helpful sofa-lifting partner, takes work. Think of an occasion when you have listened to a talk, or gotten driving directions, or just had a conversation with someone, and consider the time and cognitive effort needed in order to focus your attention on what the speaker was saying, figure out what they meant, and assess whether to accept it. Assuming that you are cooperatively disposed toward a speaker, knowing that there is something that they intend you to think can motivate you to invest the cognitive effort that might be required to figure out *what* they intend you to think. And once you have figured that out, you can invest further effort that might be needed to decide whether to think it. A speaker whose addressee is investing effort in these ways is at an enormous advantage compared to one whose addressee is ignorant of the speaker's intentions, and so who can't direct their cognitive resources in ways that complement the speaker's own effort. This is one of the ways in which human communication is at its best and most efficient when it takes the form of a cooperative joint activity.

Of course, there are limits to any communicative strategy that relies on cooperativity. Your addressee might have conflicting goals or beliefs. They might think that you are ignorant or untrustworthy. Knowing that you intend to change their mind could lead them to erect psychological defenses. In situations like these—

think of my daughter and her food aversions, or a philosophical interlocutor who won't just take your word for it about a certain piece of evidence for your view—we are forced to fall back on less efficient options like reverse psychology or providing direct evidence for the belief that we want to instill.

But the advantages of cooperating are huge. The efficiency boost that comes from interlocutors investing their effort toward a common end is one example. I have mentioned how this can benefit speakers, but as I made clear in Chapter Two, hearers also benefit enormously when speakers invests effort to customize their communicative acts with addressees' information, preferences, and goals in mind. I will discuss several other communicative advantages of conversation planning in this chapter. Indeed, I think that the communicative advantages of cooperative, joint planning are so significant that they constitute a powerful selection pressure in favor of cooperation-conducive traits, operating at both the evolutionary timescale (to shape human nature) and the cultural timescale (to shape our second natures, our conventions, and our norms). Among the traits thus selected for include the basic pro-social dispositions that allow us to coexist in large groups without (too much) conflict, the kind of memories and dispositions that allow us to engage in long-term reciprocity, mental representations of social structure and reputations that allow us to track the trustworthiness of other group members from our information about their past interactions (Planer and Sterelny, 2021), and—most importantly for my purposes—the kind of mindreading and planning capacities that allow us to coordinate and act on shared plans. I turn now to a deeper discussion of this last trait.

2 Shared Plans

What, exactly, does it mean to say that a certain episode of communication is a cooperative joint activity, governed by shared plans?

A fruitful answer to this question can be extrapolated from Michael Bratman's (1992; 2014) planning theory of shared agency. In Chapter Two, I drew on Bratman's planning theory in order to explain how individual speakers design their communicative acts for their addressees, forming communicative intentions along the way. Bratman's theory is based on the idea that many of the distinctive features of human agency can be explained by the way in which we treat our intentions as

the building blocks in complex plans. Any given intention is likely to be partial, in that it represents a goal but not a complete way of accomplishing it. In order to act on this kind of partial intention, we must flesh it out by adopting subplans that settle questions about its means and implementation. In doing so, we treat it and our other intentions, beliefs, and the pressure to maintain a coherent state of mind as constraints. By iterating this process, we construct complex, hierarchical plans that are responsive to the available information about our circumstances, and that can convert abstract goals into specific bodily movements. We are thereby able to perform more complex actions and achieve more abstract and ambitious goals than we would otherwise manage.

Bratman's theory of joint agency uses ingredients from his planning theory to develop an austere account of what it is to make shared plans and act together. The crucial ingredient is for two agents to have a shared intention, which on Bratman's view requires that they each intends some end, intends to pursue this end by means of meshing subplans, and is aware of the other's intentions.[2] To have meshing subplans of a shared intention is to have intersubjectively coherent intentions about how to accomplish it that together add up to a complete representation of a way of realizing the intention. For example, if you and I have a shared intention to play a game of table tennis, then each of our plans about our own contributions to the game must fit together. If I form an intention to stand on the north side of the table and serve first, you will need to form an intention to stand on the south side of the table and return first. It will be much easier to achieve meshing subplans of this kind, particularly when it comes to complex joint actions that involve many contributions from multiple individuals, if we know about each other's relevant

[2]Bratman requires that agents with a shared intention commonly know that they they both intend to pursue a shared end via meshing subplans (Bratman, 2014, 57–58). I think that this condition is too strong. One reason is that I doubt that we ever achieve full-blown common knowledge (Lederman, 2017). Another, which I will defend later in this book, is that the role of common knowledge and similar notions in the theory of communication has been greatly exaggerated (see also Harris 2020 and Jankovic 2014). Here is a preview of what I will say later: I think that the better our information about each other's relevant plans and expectations, and about their expectations about our plans and expectations (and so on), the better the joint-planning process will go. But there are also both cognitive and practical costs associated with every step in the direction of common knowledge, and so we trade off, doing the work to inch closer (albeit never particularly close!) when the stakes are high and cognitive resources are abundant, but settling for less when the stakes are lower or we are short on resources.

intentions and expectations. In general, we can follow Bratman's usage by saying that if A and B have a shared intention, A has a subplan of this intention, and B expects A to execute this subplan, then A's subplan and B's expectation "support" each another.

3 Conversation Plans

In the following sections, I will articulate some of the ways in which it is fruitful to theorize about human communication using Bratman's framework. It will help to start by walking through a schematic representation of what I will call a "conversation plan"—which is the sort of planning structure that I take to be at work in an extended communicative exchange. In Figure 3, I separate out several kinds of intentions that feature as elements in typical conversation plans (though not all conversation plans need to share all of these elements or have this shape).

At the top of the diagram are the interlocutors' individual intentions that they are seeking to act on in the conversation. For example, suppose we are having a conversation about where to eat dinner. I intend to eat a low-calorie meal, while you intend to carb-load for the half marathon you're running tomorrow. These intentions aren't shared, but they will figure into how the conversation plays out, because further elements in the conversation plan will have to be subplans of them, or at least be consistent with them. Next there are the various shared intentions that govern a conversation, which interlocutors form as a means of pursuing their individual intentions. For example, each of us intends that we will eat together, and that we will work out how to do this by means of a polite conversation, and we intend to pursue these ends by means of meshing subplans. These elements in the conversation plan meet the definition of shared intentions given above. Among the shared intentions in our conversation plan is our shared intention to answer what Roberts (2012b) calls the "immediate question under discussion," or "QUD." At a given stage in the conversation, the QUD might be something broad, such as, "Where will we eat dinner?" or it could be some more specific subquestion by means of which we are trying to answer that broad question, such as, "Will we eat at Veselka?" In many cases, the interlocutors' plan to answer the immediate QUD is the shared intention that is most directly governing a conversation, and so it determines which conversational moves will count as relevant. For example, if our

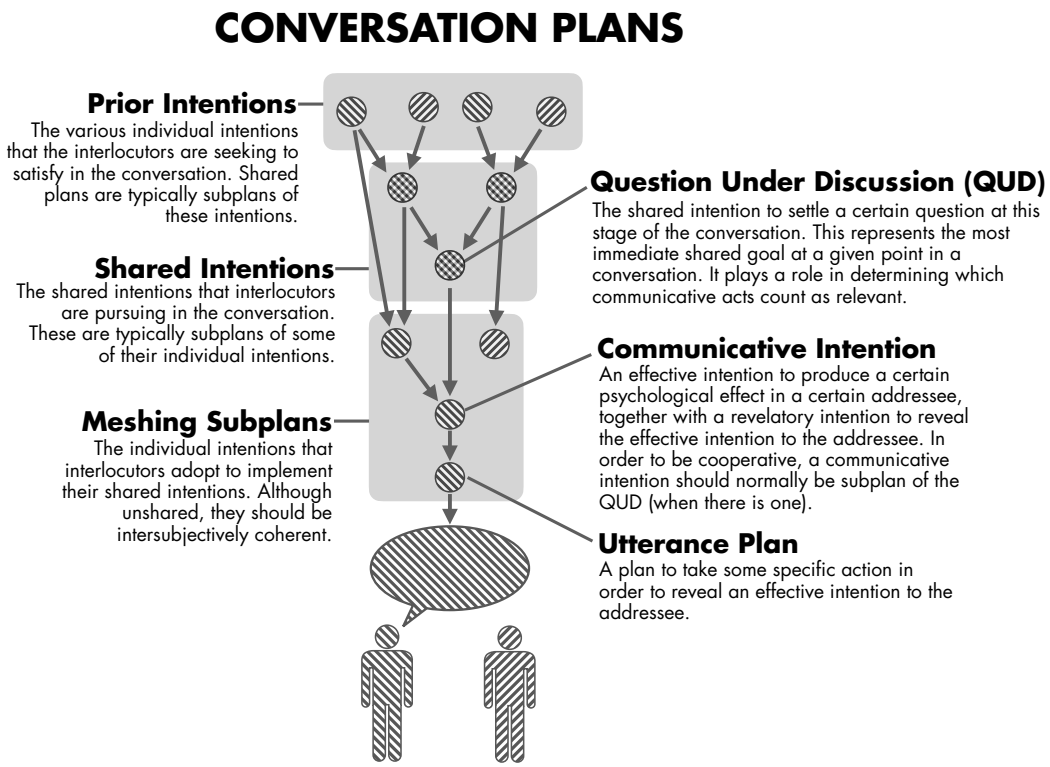

Figure 1: A schematic conversation plan. This is intended merely as an illustration of the sort of structure and elements that a conversation plan might have. Not all conversation plans need have the same structure or elements given here.

immediate conversational goal is to decide whether to go to Veselka, then asking whether it's raining will be irrelevant unless there is some obvious dependency between whether it's raining and whether we go to Veselka. (I will say much more about questions under discussion in §4, and about how we infer dependencies of this kind in §5.)

Below the level of shared plans, there are interlocutors' meshing subplans, which are their individual but (hopefully) coordinated intentions about how to accomplish their shared goals. These may include things like speakers' intentions to give the other speaker a chance to say something, the intention to pay careful attention to what the other is saying, and so on. Most importantly, there are communica-

tive intentions—the animating force behind individual communicative acts, which define their success conditions. In §4, I will argue that in the context of many cooperative conversations, communicative intentions should be suitable subplans of interlocutors' shared plan to answer the QUD. Of course, as I argued in Chapter Two, even once a communicator has a communicative intention, they must still make decisions about how to reveal their effective intention to their addressee. I have grouped together all of the intentions that result from these decisions under the label "utterance plan." This includes things like the intention to utter one word rather than another, to speak in a friendly voice, and so on.

It is important to point out I don't think that all conversation plans need to have the same structure or exactly the same elements as the schematic one depicted in Figure 3. I also don't think that every episode of human communication happens against the background of a conversation plan at all. If I honk my car's horn at a driver who has recklessly cut me off, I may succeed in getting the driver to recognize my intention to inform him that he is rude and I am angry without the benefit of any shared plan. Rather, I think of conversation plans as shared cognitive structures that, when we have them, make for much more efficient communication—and, in particular, extended communicative exchanges. I will elaborate on this point in §7.

4 Questions and plans

I have claimed that conversation plans often include a shared intention to address the immediate question under discussion (QUD), and that in these cases, communicative intentions must be subplans of this shared intention in order to be cooperative. In this section, I will substantiate these claims by reviewing the evidence for an influential inquisitive model of conversation, due to Craige Roberts (Roberts, 2012b), and then by arguing that both the successes and limitations of this model are best explained by taking it to be a model of elements in larger conversation plans.[3] Part of my argument will be that we can understand and lift the idealiza-

[3] There have been a number of other semantic and pragmatic models that take conversations to revolve around questions that participants are trying to answer. See, for example, Carlson (1982); Ginzburg (1994, 1995a,b); Groenendijk and Roelofsen (2009). I think that many of the morals that I will draw in this section could be drawn from these other models as well, but I will focus on Roberts' model for concreteness and because it has been particularly influential.

tions of Roberts' model only by interpreting it in the way that I propose. However, it will be helpful to approach that point by first introducing the idealized model itself.

Roberts' model is built as an extension of Robert Stalnaker's (1978; 2014) own idealized model of conversation. This model most cleanly applies to a specific kind of conversation, sometimes called called "communal inquiry," in which two or more interlocutors attempt to pool their information. The already-pooled information at a given stage of the conversation is the "common ground." Stalnaker has given a series of subtly different recipes for reducing common ground to the interlocutors' psychological states, most recently analyzing it as the collection of propositions that the interlocutors commonly accept for the purpose of the conversation (Stalnaker, 2014).[4] Common ground is modeled as a set of propositions, and propositions as sets of possible worlds. The "context set" is the set of worlds compatible with all of the propositions in the common ground—equivalently, the intersection of the propositions in the common ground.

Propositions can wind up in the common ground in different ways. For example, if a bell loudly rings while we are all sitting in a room, it will become common ground that the bell has wrung, even if nobody mentions it. This is because each of us heard it, could infer that the others had heard it, could infer that they could infer that we had all heard it, and so on. One way to put a proposition into the common ground, assuming that the others are paying attention and able to understand, is to assert it by uttering the right declarative sentence. By doing this, the participants can take turns building up the common ground and, *ipso facto*, shrinking the context set, thereby zeroing in on the actual world and pursuing the goal of communal inquiry.

[4] A group commonly accepts a proposition p if and only if each group member accepts p, accepts that each group member accepts p, accepts that each accepts that each accepts p, and so on, ad infinitum. Acceptance is a bit like belief, in that it involves treating a proposition as true, except that we sometimes accept a proposition for some purpose even if we don't actually believe it (Stalnaker, 1984, 79–80). For example, if I am having dinner with my wife's uncle, I might accept some of his bizarre political beliefs for the purpose of the conversation just so that I can avoid getting into an annoying debate with him. Likewise, we sometimes temporarily accept a proposition for the purpose of ultimately disproving it by a *reductio-ad-absurdum* argument. For some of Stalnaker's earlier analyses of common ground, see Stalnaker (1974; 1978; 1984; 2002). Other authors have defended alternative analyses (see, e.g., Clark (1996, ch.4)). I will discuss Stalnaker's model of conversation, and common ground in particular, in much more detail in Chapter Five.

But of course, even in the context of the sort of communal inquiries that Stalnaker's model captures best, it's not as if we proceed by asserting any old thing that will eliminate some open possibilities. Consider the following conversation:

(1) Ann: Johnny Cash once went by the nickname, 'the Undertaker.'
Bob: I am a little bit hungry.
Ann: The Hawaiian island Lānaʻi has a population of about 3300 people.
Bob: My mother's maiden name has six letters.

Here, Ann and Bob assert a series of true propositions that are (let's stipulate) not already common ground for them, thereby adding those propositions to the common ground and eliminating some hitherto live possibilities from the context set. But outside of a very unusual context, this looks like less of a conversation than a series of unrelated utterances. In real conversations, including those that most closely approximate Stalnakerian joint inquiry, interlocutors normally have specific goals about which information they are seeking, and utterances that don't further these goals are treated as irrelevant, and so infelicitous. The strangeness of (1) lies in the fact that it's very difficult to think of any conversational goal that could unify the successive utterances into a conversation, such that each of Ann's and Bob's utterances, in turn, would count as progress toward that goal.

Roberts' model is designed to capture this way in which conversations are organized around informational goals. She does this by complicating Stalnaker's model of context, adding a "question stack" alongside the common ground. The top item on the stack at a given stage of the conversation is the "immediate question under discussion," or QUD. The intuitive idea is that the immediate goal of a conversation can be represented as the question that the interlocutors are trying to answer at that moment. Whether a communicative act counts as relevant depends on whether it makes progress toward answering the QUD. An assertion does this if it at least partially answers the QUD by ruling out at least one possible answer that is compatible with the common ground. Asking a question can be relevant if its content is a subquestion of the QUD, which is to say that a complete answer to it would be at least a partial answer to the QUD. Normally, asking a question—for example, by literally uttering an unembedded interrogative clause—is intended to add a new topmost item to the question stack and thereby establish a new QUD. Following

Carlson (1982), Roberts says that to ask a question is to perform a "setup move," which focuses interlocutors' attention on a collection of answers that could constitute alternative paths toward their informational goals. Assertions, by contrast, are "payoff moves," which rule out at least some of those paths by adding information to the common ground.

Formal semanticists typically model the semantic content of a *wh*-interrogative clause as a set of propositions—intuitively, the set of its possible complete answers.[5] The QUD, like the other elements on the question stack, is itself a set of propositions. At any stage of a conversation, we can use the QUD at that stage to partition the context set at that stage into the set of *live* complete and mutually exclusive answers (i.e., those that haven't been ruled out by the common ground). This idea lends formal precision to the idea that the QUD focuses interlocutors' attention on a set of alternative possible paths forward in the conversation. Once such a partition is set up, the interlocutors sets to work on eliminating cells, thereby homing in on an answer. This intuitive idea can be turned into precise predictions about which assertions and questions are relevant. An assertion is relevant when its content is incompatible with at least one cell in the partition—i.e., if it at least partially answers the QUD by ruling out one of the still-live answers. Asking a question q is relevant if q is a subquestion of the QUD, which is to say that a complete answer to q would be a partial answer to the QUD. Roberts and others have showed that these predictions are empirically borne out by our relevance judgments about a wide range of discourses of the following kind:

(2) S: What did Hilary eat?
A1: She ate tofu.
A2: She went swimming.

(3) S: What did Hilary eat?
A1: Did she eat tofu?
A2: Did she go swimming?

We can explain our judgments about these cases by saying that S's question installs a new QUD, to which A1's response, but not A2's, is relevant.

[5] This is a simplified amalgamation of ideas first proposed by Hamblin (1973), Groenendijk and Stokhof (1984), and von Stechow (1991).

Inquisitive models have been used to explain a range of other phenomena, all rooted in the idea that we use cues about relevance to understand others and to make ourselves understood.[6] Roberts' own other case study has to do with the sort of intonation patterns that linguists call "prosodic focus." Building on work by von Stechow (1991), Roberts shows that which sorts of prosodic focus are felicitous is itself a function of the QUD.

(4) S: Who invited Sue?
A1: $[\text{Mary}]_{\text{F}}$ invited her.
A2: Mary invited $[\text{her}]_{\text{F}}$.

(5) S: Who invited whom?
A1: Who did $[\text{Mary}]_{\text{F}}$ invite?
A2: Who did Mary $[\text{invite}]_{\text{F}}$?

In exchanges like these, S's question establishes a new QUD. Roughly speaking, responses to the question that include prosodic focus are felicitous only if they would have the QUD as their content if the focused element (here marked with $[]_{\text{F}}$) were replaced by a *wh*-phrase. Even more roughly, this is to say that an answer is congruent with a question if the focused element in the answer corresponds to the *wh*-phrase in the interrogative clause used to ask the question.[7] These turn out to be very robust generalizations about when it is appropriate to use prosodic focus. Roberts explains this phenomenon by arguing that the communicative function of prosodic focus is to probe the QUD as a way of reinforcing the interlocutors' attention on their shared conversational goal of the moment.

As the above examples illustrate, asking a question is one way to install a new QUD at the top of the question stack, and making an assertion is one way of removing a question—i.e., by answering it. The question stack is a *stack* in the sense that when a question is added, it pushes down the previous question to the next spot in the stack, and when the QUD is resolved, it leaves the stack, returning the next unanswered question down to the top. This aspect of the model is designed

[6] See, *inter alia*, work on projective content (Simons et al., 2017, 2010), loose talk and metaphor (Hoek, 2018), disjunction (Simons, 2001), epistemic modals (Beddor and Egan, 2018), attitude verbs (Schaffer, 2007; Yalcin, 2018), and our ability to interpret semantically underspecified expressions in general (Schoubye and Stokke, 2016).

[7] For the more technically precise version of these ideas, see Roberts (2012b, §2.1).

to make sense of the way in which we sometimes pursue an inquiry by engaging in temporary subinquiry, then returning to the main thread after the subinquiry has been resolved. For example, suppose we're trying to decide where to go to dinner. The QUD might be, "where should we have dinner?" As a strategy for eventually resolving this question, we temporarily turn to the question, "Which of us feels like Mexican food?", any complete answer to which, together with the common-ground assumption that we should eat something that everyone wants, will eliminate at least some of the live answers to the original QUD.

A new QUD can also get established via nonlinguistic channels, just as information can sometimes make it into the common ground without anyone asserting anything. Just as a bell ringing might make it common ground that a bell has wrung, the same event might sometimes also establish the question, "who rang that bell?" as the QUD. This is why it makes sense to say things like the following after a conspicuously ringing bell:

(6) $[\text{Steve}]_{\text{F}}$ must have wrung it.

(7) Did $[\text{you}]_{\text{F}}$ ring that bell?

Of course, the participants might be in some doubt about whether this has become the new QUD until someone utters one of these sentences, but this gives speakers all the more reason to use prosodic focus in these ways, thereby reinforcing what may have been somewhat indeterminately the QUD already.

This concludes my summary of the basics of inquisitive models. On to my main order of business: What is Roberts' model a model *of*? What is its subject matter? What facts about real conversations are there in virtue of which the model generates all of its useful predictions and gives an at least somewhat accurate representation of the structure of discourse?

My answer to this question is that QUDs are best understood as a models of interlocutors' plans to answer questions. In conversations that conform best to Roberts' model, what makes a certain question the QUD of the conversation is that the interlocutors have a shared intention to answer it, and this is the farthest-downstream shared intention in their conversation plan. I will call this shared intention to answer the QUD the "QUD-plan." This way of thinking about QUDs explains the successes of Roberts' model, but also the nature of its idealizations.

This interpretation of the model is consistent with Roberts' own remarks about her subject matter. She introduces her theory by telling us that "[d]iscourse is organized around a series of conversational goals and the plans or strategies which conversational participants develop to achieve them" (Roberts, 2012b, 3). She also sometimes refers to QUDs as "discourse goals" (Roberts, 2012b, 26), distinguishing these from "domain goals," which are "things we want to achieve quite apart from inquiry" (Roberts, 2012b, 7). But in later refinements of her model, Roberts describes discourse goals as a "distinguished type of domain goal, those the interlocutors are jointly committed to achieving in the discourse itself" (Roberts, 2018, 323).[8]

So why think that QUD models are models of elements in conversation plans? The first and most important reason is that this interpretation gives an independently motivated explanation of how a QUD plays its central functional role in determining which conversational contributions are relevant. For a communicative act to be relevant to the QUD, on this view, is just for the speaker's communicative intention to be a coherent and constructive subplan of the interlocutors' shared intention to answer the QUD. Suppose that you and I have a shared intention to answer the question, *What did Hilary eat?* In this context, if you know that she ate tofu, it makes sense for you to pursue our common goal by forming an effective intention to convey this information to me.[9] In order to ensure that your effective

[8] Although Roberts does not directly cite Bratman as an influence on her theory, she does repeatedly mention "Planning Theorists in artificial intelligence," citing a number of authors who themselves drew on Bratman's theory of intentions and practical reasoning in building computational models of discourse. E.g. Cohen and Levesque (1990); Cohen and Perrault (1979); Grosz and Sidner (1986); Perrault (1990); Thomason (1990). In the Afterward to her original paper, Roberts explains that she was influenced by Bratman's work while she was a postdoc at Stanford in 1986–1988, where he was developing his theory at the time (Roberts, 2012a, 3).

[9] It might be objected that in order to settle the QUD in a way that truly fits with the model, a speaker should form an intention to add their answer to the common ground, and not just inform their interlocutor(s). I will ignore this detail for now but I'll return to it in great detail in Chapter Five, where I'll argue, *inter alia*, that communicative intentions are normally directed at changing addressees' mental states rather than the common ground. Here's the short version of what I will say there: There are many contexts in which we can't make information common ground, but in which we still manage to achieve our communicative goals just fine, and this is because what we care about in those contexts is changing our addressees' private states of mind. And in contexts where we do seek to add p to the common ground (insofar as this is *ever* possible, which it might not be), a crucial means to this end would be to get our addressee to believe p (or at least accept p for the

intention meshes with my own subplans of the QUD-plan, you should also reveal it to me. If you form these intentions, you *ipso facto* have the sort of communicative intention required to assert that Hilary ate tofu. The fact that your communicative intention is a constructive and coherent subplan of the QUD-plan is what explains our judgment that your assertion would be a relevant one. By contrast, if you were to form a communicative intention to inform me that Mary went swimming, this would *not* be a constructive subplan of the QUD-plan, because it wouldn't be a good way of answering the QUD. This explains our judgment that *this* assertion would be *ir*relevant.

We can explain relevance judgments about questions in a similar way. In the same context just described, in which the QUD is, *What did Mary eat?*, I might decide to pursue our shared QUD-plan by means of proposing a new subplan—in this case, a shared plan to answer the subquestion, *Did Mary eat tofu?* Again, in order to make sure that you our contributions to the plan mesh, I should also try to reveal this intention to you. Now I have the communicative intention that one needs to have in order to perform a communicative act of asking whether Mary ate tofu. This act will seem relevant insofar as the pursuit of the plan that it proposes is a good means to our prior end of answering the QUD, which it is, since my new question is a subquestion of the QUD. By contrast, if I were to ask whether she went swimming, I would be proposing a plan that would not be a good subplan of our QUD-plan, and so we would judge that question to be irrelevant.

On this view, questions reside in the question stack by virtue of interlocutors' shared plans to answer them. Contiguous questions in the stack must stand in subquestion relations because they represent intentions that must stand in subplan relations, and we can, in general, make progress toward answering a question by answering its subquestions. The question stack is a formal model of a hierarchical plan made up of shared intentions, which together make up one substructure within a conversation plan.

There are a number of significant ways in which Roberts' model, as well as the explanation that I have given of its subject matter, is idealized. Many conversations are much less well behaved than the foregoing story suggests, or lack some of the features that, I have suggested, explain the model's success. Some of these idealizations are inherited from Stalnaker's own common-ground-based model, and I will

purpose of the conversation). See Harris (2020) for a more detailed version of this argument.

address them in Chapter Five. But some of the idealizations are specific to QUD models themselves. I will address these here. Ultimately, I think that understanding these idealizations and formulating a strategy for lifting them will lend further support to my interpretation of the model and my explanation for its success.

First, consider the fact that when we ask a question, we often care about which of our interlocutors answers it. For example, suppose I am having a conversation about where to eat with several colleagues, most of whose taste in restaurants I don't trust. So I turn to the one colleague I trust most, make eye contact with them, and say,

(8) Where do you want to eat?

In this case, I am not proposing a shared plan to answer the question that just any of my interlocutors should feel free to act on. I intend for my foodie colleague, specifically, to answer. If one of my other colleagues now makes a proposal about where to eat, they will be speaking out of turn. It therefore seems wrong to think of my question as aiming to establish a QUD that models a shared intention to which we all stand in a symmetrical relationship. Call questions of this kind *directed* questions. I think that directed questions are quite common—probably more common than undirected questions. In one-on-one conversations, a questioner typically wants their addressee to answer, and it would be strange to ask questions and then answer them oneself. In larger conversations, it's normal to direct questions, either using vocatives ("Steve, where should we eat?") or using eye contact or some other nonverbal signal.

By taking the function of acts of questioning to be the establishment of new QUDs, and by thinking of the QUD as representing a shared plan, the model that I have presented idealizes away from directed questions. We could complicate the model to make sense of them—for example, by positing a separate question stack for each interlocutor and thinking of undirected questions as a special case in which a question gets added to all interlocutors' stacks at the same time. But this new model would also require a new interpretation, to explain what it would be for a question to reside on one interlocutor's stack but not others'.

If we think in terms of the larger conversation plans of which shared QUD-plans are just one part, I think a clear interpretation recommends itself. In some cases, the interlocutors in a conversation take up a shared intention to resolve a

question, but in other cases the intention to resolve a question belongs to just one interlocutor. When I turn to my foodie colleague and utter (8), I intend for them—just that one colleague—to form an intention to answer my question. If one of our other interlocutors interprets me as intending them to form a similar intention, then they have misunderstood my communicative act. Of course, the intention that I am trying to produce is still a subplan of our shared plan to choose a place to eat. But by directing my question at one interlocutor, I am, in effect, proposing a division of labor in how to accomplish that shared plan, tasking just them with the responsibility to make make a suggestion. The right way for our other interlocutors to respond to my question, insofar as they don't object to this way of proceeding, is to form whatever intentions and expectations are required in order for their own contributions to the conversation plan to mesh with the intention that I am proposing for my foodie colleague. In this case, that might entail an expectation that the foodie will answer, together with an intention to wait their turn and pay attention to this response.

If this line of thought is right, then a question is not always the QUD of a conversation by virtue of our shared plan to answer it, but by virtue of one interlocutor's intention to answer, together with other interlocutors' meshing expectations and intentions. But this understanding of QUDs is available to us only insofar as we take them to be models of components in larger conversation plans. It just turns out that which kinds of components they model can vary from one case to another.

Another way in which the QUD model is idealized is that it ignores the possibility of felicitous changes of topic. This follows from what Roberts says about the options for removing items from the stack—a position that she explains in terms of the fact that QUDs represent interlocutors' plans:

> If a question is accepted by the interlocutors, this commits them to a common goal, finding the answer; like the commitment to a goal in Planning Theory, this is a particularly strong type of commitment, one which persists until the goal is satisfied or shown to be unsatisfiable. (Roberts, 2012b, 5)

Some conversations really do work like this. If we are talking about whether a certain restaurant is good because we urgently need to eat—if we're all starving, say, and we only have a half hour to spare for lunch—then it will be genuinely norm

violating for you to start talking about something else before we've resolved the question (along with the superquestion of where to have lunch). But not all conversations are like that. Imagine another, more casual and meandering conversation with friends in which we float from topic to topic relatively freely. In this conversation, we might momentarily alight on the question of whether that same certain restaurant is good, and then, without resolving it, wander off to another topic without any infelicity. Many of our best conversations exhibit this sort of much looser flow.

Roberts' model lacks the resources to explain the difference between these two kinds of conversation. At the relevant moments of both conversations, the QUD is the same. The difference, I submit, can be found higher up in the conversation plans that govern these conversations. In the former conversation, our discussion of whether the restaurant is good is happening in the service of our shared intention to eat lunch as soon as possible. This intention is incompatible with our having a meandering chat about other topics at the moment, and so any intention to talk about something else will lead us to have an incoherent conversation plan. By contrast, the more meandering conversation is the product of less urgent, more exploratory goals. Our purpose is not to meet any immediate material need, but to enjoy each other's company, entertain ourselves, and maybe acquire some new information along the way insofar as that's compatible with having fun. If our QUD-plan is a subplan of shared intentions like these, then it doesn't matter if we abandon one QUD in favor of another. If we have wrung most of the enjoyment out of the first topic, or if we think of another thing that will be more fun to talk about, then changing topic will better serve our ultimate conversational goals.

Meandering conversations are similar to other activities in which features of our upstream plans allow us to adopt a flexible attitude toward our downstream intentions. When I play with my young children, we tend to cycle through many different activities, sticking to each one only as long as it holds their sometimes-fleeting attention. We might begin drawing a picture, then get distracted by a a tickle fight, then take a break for a snack, and then (perhaps before the snack is finished) listen to some music. It is quite common that each activity goes unfinished before we move on to the next thing. But that's okay, because the overarching aim is to entertain and stimulate my kids, and so there may be no need to follow through on our subplan to draw the picture if doing that is no longer fun. By contrast, if our

intention to draw a picture were the subplan of our intention to complete a class project that was due the next day, things might be different. In general, we sometimes pursue the same means to different ends, and features of the ends can shape how, and how urgently, we pursue the means. The same goes for conversations, in which the manner and urgency of our attempt to answer a certain question can be shaped by our upstream reasons for asking it in the first place.

Roberts' model idealizes away from these considerations, and ignores meandering conversations at the expense of giving us a better model of more strictly organized ones. This is no objection to the model; that's just the sort of thing that idealized models do. But if we ever want to lift this idealization, and to come up with a model that can do more justice to this way in which conversations can vary, it will be important to think of QUDs as models of intentions that are themselves elements in larger conversation plans, and to design the successor model to capture more of the elements in these plans. The same holds if we would like to lift the idealization that prevents Roberts' model from making sense of directed questions. Again, in raising this limitation of the model, my aim is not to criticize it. Rather, my point is that reflecting on these limitations of Roberts' model can tell us something important about what it is a model *of*. The real subject matter of Roberts' model are conversation plans, which can include shared plans to answer questions, but other kinds of intentions that may lie either upstream or downstream from these shared plans.

5 Cooperativity and Practical Rationality

Consider again example (2), from the last section:

(2) S: What did Hilary eat?
A1: She ate tofu.
A2: She went swimming.

Rather than simply label A2's response as irrelevant, we might be tempted to think that A2 meant something that was relevant but indirect. For example, perhaps A2 was implicating that Hilary swam instead of eating, or couldn't have eaten because you're not supposed to swim after eating. Roberts' model gives us a partial explanation of how we are able to recognize relevance implicatures of this kind (and,

therefore, of why it is reasonable for speakers to communicatively intend to convey them). An interpreter could reason as follows:

(9) Calculating the implicature in (2):

(i) Being cooperative requires saying relevant things;

(ii) asserting something relevant requires partially answering the QUD;

(iii) the literal content of A2's utterance is not relevant by this standard;

(iv) the assumption that A2 is being cooperative therefore requires that they must indirectly mean something else;

(v) whatever this indirectly meant proposition is, it has to entail a partial answer to the QUD.

This reasoning falls short of being an algorithm for identifying the implicature, since it doesn't tell us precisely which proposition the speaker indirectly meant. But it does give a significantly better explanation of how we calculate relevance implicatures than Grice was able to offer, because it places severe and relatively specific constraints on what a cooperative speaker could have meant, thereby narrowing the interpreter's search space. The QUD plays a crucial role, both at the stage of recognizing that there is an implicature and at the stage of narrowing down what its content could be.

The idea that the QUD is a model of a shared plan can help us to make sense of how (9) works. When a speaker implicates or otherwise indirectly means some proposition p, p is normally only loosely related to the linguistic meaning of the speaker's utterance. The interpreter therefore needs some fairly rich source(s) of extralinguistic evidence in order to infer the speaker's effective intention. This evidence could be just about anything. For example, in (2), it would be useful to know whether the speaker (or Hilary) has particularly strong beliefs about whether eating before swimming causes cramps. The inference in (9) is based on the idea that one particularly common and useful kind of evidence about speakers' effective intentions is information about the rest of the conversation plan of which their effective intention is a subplan. In general, if we know about the ends that a person is trying to achieve, and if we assume that they are well informed and rational, this can help us to infer their intended means to those ends. For example, if I think that you are

rational, that you intend to buy groceries at 11pm, and that you know that only one grocery store is open at that time, then it will be reasonable for me to guess that you will form an intention to go to that store. The inference in (9) is an instance of this kind of inference from knowledge of a person's ends to a guess about the means that they might take to that end: S infers that whatever A2 communicatively intends, this intention must be a coherent and constructive subplan of their intention to answer the QUD, and this conclusion serves as rich evidence about the content of A2's effective intention.

This fits with how Roberts understands her own concept of relevance. In introducing it, she emphasized that relevance is always relevance *to a given QUD at a given time*, and explicitly compared this to "Grice's relativization of his Maxim of Relation to 'the purposes of the discussion' " (Roberts, 2012b, 21).[10] This comparison makes sense if to say something that is relevant to the QUD just is to do something that furthers the purpose (i.e. *the plan*) of the conversation.

This explanation generalizes beyond relevance implicatures. Consider one of Grice's examples of quantity implicature—a reference letter that says only that "Mr. X's command of English is excellent, and his attendance at tutorials has been regular" (Grice, 1975). How does the reader of this letter infer the writer's intention to disparage Mr. X's philosophical ability? An important clue, I submit, is that the purpose of this "conversation" (and of reference letters in general) is for the writer to convey as many positive things about the applicant's relevant abilities as they can.[11] If the writer is being cooperative with respect to this shared aim, they must intend us to infer that they have nothing more positive to say. But suppose

[10] Somewhat ironically, the maxim of relation is one of the maxims that Grice did *not* explicitly relativize to the purpose of the conversation in this way. Nevertheless, I think he should have, and it is plausible to think that he took this relativization to be so obvious as to not be worth spelling out. See below.

[11] It might be objected that it is strange to say that the writer and reader of a reference letter have a shared conversation plan, given the spatiotemporal distance between writing and reading, and since the writer may not even know who the reader is. I don't think this is right. The writer and reader of the letter each intends that they engage in the transmission of information about the applicant's abilities, and intend to play interlocking roles in the transmission of this information via meshing subplans (e.g., the writer's plan to say various things and the reader's plan to pay attention to what the writer says, read between the lines, etc). There is nothing in the theory of shared plans that I sketched in §2 that rules out shared plans of this kind. Thanks to Hans Kamp for pressing me on this point.

we vary the conversational purpose by imagining that the author is writing a different kind of letter—one whose generally accepted function is merely to confirm Mr. X's English-language and time-management capacities. Now the implicature disappears, because there is no threat of incoherence between conversation plan and literal meaning to trigger our search for it. This is why the maxim of quantity tells us to make our contributions "as informative as is required (for the current purposes of the exchange)" (Grice, 1975, 45).

Grice didn't mention conversational purposes in his formulations of the other maxims, but he should have. What counts as following or flouting each of them depends on interlocutors' shared plans in just the same way. We have already seen this in the case of the maxim of relation, which Grice formulates simply as "Be relevant," but which Roberts upgrades with her relativization to the QUD.[12] Similar considerations apply to the maxim of manner. Take one of Grice's examples of a manner implicature, wherein a reviewer could have said "Miss X Sang 'Home sweet home' " but instead said, "Miss X produced a series of sounds that corresponded closely with the score of 'Home sweet home'." In this case, we are tempted to conclude that the reviewer intended to disparage Miss X's singing. Grice explains this by saying that the best explanation for why the reviewer "selected the rigmarole in place of the concise and nearly synonymous *sang*" is that they intended "to indicate some striking difference between Miss X's performance and those to which the word *singing* is usually applied" (Grice, 1975, 56). This is a plausible, if incomplete explanation, but notice that it depends on the assumption that the participants in this conversation are seeking to efficiently convey information. Imagine, instead, that the interlocutors have agreed that the speaker will review Miss X's singing while also doing an impression of a notoriously prolix rival, and the implicature disappears, because there is no longer any threat of incoherence between the conversation plan and the speaker's utterance. Just what it takes to "avoid unnecessary prolixity" (Grice, 1975, 46) depends on what sort of utterance is necessary in order to further the conversation plan.

[12] Notably, Grice recognized the weaknesses of his own formulation: "Though the maxim itself is terse, its formulation conceals a number of problems that exercise me a good deal: questions about what different kinds and focuses of relevance there may be, how these shift in the course of a talk exchange, how to allow for the fact that subjects of conversation are legitimately changed, and so on. I find the treatment of such questions exceedingly difficult, and I hope to revert to them in later work" (Grice, 1975, 46).

Even the maxim of quality requires relativization. Or, rather, in Grice's own formulation, it already is relativized to a certain kind of conversational purpose:

> 1. Do not say what you believe to be false.
> 2. Do not say that for which you lack adequate evidence. (Grice, 1975, 46)

Of course, we can see that what counts as "adequate evidence" will itself depend on the conversational purpose. Evidence that is adequate for a casual conversation might not be adequate for a deposition. But there is also a much more general point that the maxim of quality, thus formulated, only makes sense for conversations whose aim is to exchange knowledge (or well supported beliefs).[13] Not all conversations have this aim. Grice recognized this limitation of the maxims:

> The conversational maxims, however, and the conversational implicatures connected with them, are specially connected (I hope) with the particular purposes that talk (and so, talk exchange) is adapted to serve and is primarily employed to serve. I have stated my maxims as if this purpose were a maximally effective exchange of information; this specification is, of course, too narrow, and the scheme needs to be generalized to allow for such general purposes as influencing or directing the actions of others. (Grice, 1975, 47)

There are a number of conversational aims other than influencing or directing actions that don't sit well with the maxim of quality as it is articulated by Grice. For example, sometimes our overarching goal is to joke around, to test out ideas to which we don't yet want to commit ourselves, to tell fictional stories, and to engage in pretend play. And this matters to whether and how we interpret implicatures. In the context of a factual conversation, an utterance of "You are Sherlock Holmes" is likely to be understood as conveying an indirect meaning about the addressee's sleuthiness, but it might carry no such implicature in the context of a game of Victorian-England pretend play, in which the utterance poses no risk of incoherence with the conversation plan.

[13] The same point applies to Grice's formulation of the maxim of quantity.

Notice that whereas the relevance calculation given in (9) makes reference to the QUD, the analogous calculations that would be needed to derive the manner and quality implicatures I've discussed would have to make reference to higher-level features of the conversation plan—superplans of the shared intention to answer the QUD. The intention to engage in pretend play is a more abstract goal than the QUD about who will be Sherlock Holmes, and the question of who is Sherlock Holmes can be asked in the service of either factual or pretend-play goals. Likewise, someone may be answering a certain QUD in a serious way or with an added intention to impersonate someone else in the act. The latter intention merely adds a constraint on how they will answer the question—a constraint that is exogenous to the QUD itself, imposed from elsewhere in the conversation plan.

Finally, consider the cooperative principle itself, which Grice clearly relativizes to the conversational purpose:

> We might then formulate a rough general principle which participants will be expected (*ceteris paribus*) to observe, namely: Make your conversational 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. One might label this the Cooperative Principle. (Grice, 1975, 44)

Grice goes on to suggest that even this is itself a special case of a broader principle; "one of [his] avowed aims is to see talking as a special case or variety of purposive, indeed rational, behavior" (Grice, 1975, 47). Although Grice gives examples of how nonlinguistic versions of the maxims would apply, he does not attempt to formulate a more general, not-necessarily-linguistic version of the cooperative principle. How might it go? I have a proposal: This general version of the cooperative principle would merely entreat agents to be practically and cooperatively rational in approximately the sense spelled out by Bratman. They should avoid incoherent combinations of intentions and beliefs, they should form intentions to do what they take to be the necessary means to their intended ends, and they should try to form meshing subplans of whatever shared intentions they have. Indeed, Grice comes very close to saying this himself, although he lacks some of the very useful theoretical affordances that Bratman later concocted:

……talk exchanges seemed to me to exhibit, characteristically, certain features that jointly distinguish cooperative transactions:

1. The participants have some common immediate aim, like getting a car mended; their ultimate aims may, of course, be independent and even in conflict-each may want to get the car mended in order to drive off, leaving the other stranded. In characteristic talk exchanges, there is a common aim even if, as in an over-the-wall chat, it is a second-order one, namely, that each party should, for the time being, identify himself with the transitory conversational interests of the other.
2. The contributions of the participants should be dovetailed, mutually dependent.
3. There is some sort of understanding (which may be explicit but which is often tacit) that, other things being equal, the transaction should continue in appropriate style unless both parties are agreeable that it should terminate. You do not just shove off or start doing something else. (Grice, 1975, 49)

On this view, the cooperative principle is merely the principle that agents involved in joint planning shouldtry to be structurally rational.[14]

If this is right, then Grice's theory of implicature is itself a partial theory of conversation plans—a theory of how we make use of them, and of the assumption that our interlocutors are making use of them, to convey information beyond what we explicitly say.

6 Joint plans and contextual parameters

[This section is still just a preview. Eventually it will include a version of the following argument: There is a widespread pattern according to which it is often acceptable for contextual parameters to evolve toward greater strictness over the course of a conversation (See, e.g., work on vagueness, loose talk, epistemic standards, etc.) The best explanation of this fact is that these contextual parameters are models of

[14]Thanks to Sam Berstler and Elmar Unnsteinsson for discussion on this point.

our shared plans about how to use words, which (like other plans) are subject to diachronic pressure to remain stable once we have committed to them.]

Shared conversation plans can also play an important role in allowing us to deal with context-sensitive language. As one example, consider how we use gradable adjectives, such as 'tall.' Suppose that Mike and Reggie are looking at a group of basketball players, each of whom is over 6 feet, 5 inches tall, when they have the following exchange:

(10) Mike: Now, *they* are tall.
Reggie: Is your cousin tall?
Mike: No.

The most influential theories of gradable adjectives tell us that the literal content of Mike's first utterance is that the basketball players are taller than d, where d is the contextually relevant degree of tallness (Kennedy, 2007). In a situation in which Mike and Reggie already agreed on a standard for tallness, and in which Mike but not Reggie could see how tall the basketball players were, Mike could have used the same utterance *factually*, to inform Reggie of their height. That's what's happening in Mike's second utterance, in which he informs Reggie that his cousin is not tall by the standard on which they have coordinated. But when Mike makes his first utterance, both he and Reggie can see for themselves how tall the basketball players are. In this situation, Mike's utterance has a *metalingusitic* aim, which is to eliminate some of the possible values of d—a contextual standard for what they will count as tall in this conversation. It is their coordination on that new standard that allows Mike to subsequently convey some information to Reggie about his cousin's height.

A number of philosophers and linguists have argued that we often use context-sensitive expressions to make metalinguistic proposals of this kind (Barker, 2002; Ludlow, 2014; MacFarlane, 2016, 2020; Plunkett and Sundell, 2013). Chris Barker (2002) thinks of these proposals as updating interlocutors' shared beliefs about the conversation's context. But John MacFarlane (2016; 2020) points out that it makes more sense to think of utterances like Mike's first as proposing shared *plans* about how to use context-sensitive expressions for the foreseeable future of the conversation. After all, there is no independent fact about which height counts as tall, over and above what the participants in a conversation decide, and so there is nothing for them to have metalinguistic beliefs about in this case other than the results of

their own decisions—i.e., their plans.[15] If this is right, then we should understand the conversation in (10) as one in which Mike and Reggie first adopt a shared metalinguistic plan, and then rely on this shared plan to make their later utterances intelligible to each other. This gives them a mechanism for packing more information into fewer, shorter utterances, but it would seem to require that they actively track each other's intentions and expectations with respect to the conversation.

7 Planning and Pragmatic Competence

I have been arguing that we can usefully think of many conversations as governed by conversation plans, which are made up of shared intentions together with interlocutors' meshing subplans of those shared intentions. This view, I have argued, can help us to understand why we form revelatory intentions—namely, because doing so allows us to recruit our addressees into the shared project of communicating with us. The idea that conversations are cooperative joint activities also makes sense of and unifies a collection of successful models in pragmatics, including Roberts (2012b) QUD-based model, MacFarlane's (2016; 2020) plan-expressivist theory of context-sensitivity, and Grice's (1975) theory of conversational implicature.

The most direct assault on this way of thinking comes from within the Gricean house, courtesy of Ray Buchanan and Henry Schiller (2022). They argue that because what a speaker means by their utterance is wholly fixed by their communicative intention, there is no role in a Gricean theory of communication to be played by principles, maxims, shared plans, relevance, or questions-under-discussion. On one hand, there is no role for these things to play in making it the case that a speaker says or means what they do by an utterance, since that role is wholly played by the speaker's intentions. On the other hand, there is no essential role for these things in the interpreter's process of recognizing the speaker's intention, because this recognition process is just an application of our general capacity for inference to the best explanation, and in general there are no principles that *must* be followed in such an

[15] Building on a formal model first developed by Gibbard (2003), MacFarlane goes on to develop a rich, "plan-expressivist" theory of gradable adjectives and other vague expressions, on which our uses of them typically function as factual–metalinguistic hybrids, both conveying information about the world and proposing more specific plans about which standards to adopt.

inference. When it comes to inference to the best explanation, there are no rules about which evidence, principles, or thought processes must be involved; anything goes.

> It would be absurd to suggest, for example, that an agent can only come to know via IBE that her friend is in pain on the basis of her yelping and recoiling her hand from the stove if there was some way of deriving that conclusion from general principles linking pain-behavior and pain; so too we claim for IBE-driven utterance intepreration. IBE simply is not a species of deductive inference from hidden/suppressed principles. (Buchanan and Schiller, 2022, 69)

From these arguments, Buchanan and Schiller conclude that no systematic theory of the kind that I have been sketching in this chapter is possible.

> Reflecting on the defeasible, and variable, role that features such as relevance play in utterance interpretation should make us suspicious of the search for robust, explanatory, predictive principles that offer much by way of improvement on the Gricean platitudes concerning what, as Grice himself once put it, 'a decent chap should be expected to do'. (Buchanan and Schiller, 2022, 77)

Buchanan and Schiller would presumably say the same about conversation plans. They have no role to play in fixing what speakers mean, and they have no essential role to play in the process by which interpreters identify what speakers mean. We can formulate no principles about how they *have to* work. Therefore, there is no point in trying to give a systematic theory about them.

I agree with all of this except the conclusion. In particular, I think that it's quite possible for people to communicate in the complete absence of a conversation plan. Imagine walking down the street, minding your own business, when a complete stranger leans out of their third-story window to shout a profanity-laced insult at you. They might succeed in getting you to recognize their intention, thereby successfully communicating with you, but you share no conversation plan with them. All that truly matters for communication, after all, is that they succeed in revealing their effective intention to you, and they might be able to do that without any of the bells and whistles that I have discussed in this chapter.

Fair enough, but that doesn't mean that we never make use of the bells and whistles. Indeed, it may be that the bells and whistles are often very important, and help to explain why human communication is so often both ambitious and successful. Concluding that conversation plans are explanatorily inert from the fact that we can sometimes succeed at communicating without one would, I think, be like concluding that power tools are pointless or inexplicable from the fact that people sometimes build houses using only hammers and hand saws. Of course the power tools aren't necessary, strictly speaking, but they make the job easier and more efficient in a reliable and explainable way, and it is perfectly possible to say something systematic about why.

In the same way, we can think of conversation plans as recurring patterns of mental representations that we regularly use to organize and facilitate our conversations, and that make these conversations much more efficient when we do. Because we can usually assume that a cooperative interlocutor's communicative intention will cohere with and flesh out the conversation plan, a detailed conversation plan gives us a powerful way of anticipating what speakers are likely to communicatively intend. This can make recognizing each other's intentions easier, but it can also allow make it rational for speakers to try to reveal more complex intentions with less linguistic evidence. None of this is a prerequisite for intention recognition, but it makes intention recognition so much more efficient that we take advantage of it whenever possible.

All of this adds up to an argument that a capacity for shared planning is an important part of what it is to be a pragmatically competent language user.[16] Moreover, as Roberts' theory of prosodic focus makes clear, natural languages themselves include grammatical systems that are designed to manipulate conversation plans. This suggests that our capacity for shared planning is required for linguistic competence itself, if we construe that concept as including whatever is required in order to use natural language in accordance with its proper function.

[16]For related ideas of "pragmatic competence," see Roberts (2017, 435; 2022), and Unnsteinsson (2022).

8 Looking Forward

The last two chapters have together illustrated the ways in which human communication is made possible by our capacity for planning. This includes both the individual planning that allows individuals to design the communicative acts, which necessitates the formation of effective intentions. It also includes the joint planning that allows us to organize our conversations, which gives us powerful reasons to form revelatory intentions. Since these two kinds of planning enable a number of features of human communication that make it significantly more efficient, I hope to have explained something about why humans would go to the trouble of communicating by means of intention recognition at all.

There is an obvious, as-yet unanswered objection to the line of thought that I have developed in these two chapters. I have argued that human communicators routinely engage in intelligent but unconscious practical reasoning and mindreading on a scale that will probably strike many readers as implausible. Do we really do all of this reasoning? Isn't there some simpler, less cognitively demanding option? I will turn to this question in Chapter Four.

References

Barker, C. (2002). The dynamics of vagueness. *Linguistics and Philosophy*, 25:1–36.

Beddor, R. and Egan, A. (2018). Might to better: Flexible relativism and the qud. *Semantics and Pragmatics*, 11(7):1–43.

Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review*, 101(2):327–341.

Bratman, M. (2014). *Shared Agency: A Planning Theory of Acting Together*. Oxford University Press, Oxford.

Buchanan, R. and Schiller, H. I. (2022). Pragmatic particularism. *Philosophy and Phenomenological Research*, 105(1):62–78.

Carlson, L. (1982). *Dialogue Games: An Approach to Discourse Analysis*. Number 17 in Synthese Language Library. D. Reidel, Dordrecht, The Netherlands.

Clark, H. H. (1996). *Using Language*. Cambridge University Press, Cambridge.

Cohen, P. R. and Levesque, H. J. (1985). Speech acts and rationality. In *23rd Annual Meeting of the Association for Computational Linguistics*, pages 49–60. Association for Computer Linguistics, Chicago.

Cohen, P. R. and Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence*, 42:213–261.

Cohen, P. R. and Perrault, C. R. (1979). Elements of a plan-based theory of speech acts. *Cognitive Science*, 3:177–212.

Gibbard, A. (2003). *Thinking How to Live*. Harvard University Press.

Ginzburg, J. (1994). An update semantics for dialogue. In *First International Workshop on Computational Semantics (IWCS 1)*, pages 111–120.

Ginzburg, J. (1995a). Resolving questions, i. *Linguistics and Philosophy*, 18(5):459–527.

Ginzburg, J. (1995b). Resolving questions, ii. *Linguistics and Philosophy*, 18(6):567–609.

Ginzburg, J. (2012). *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford.

Grice, H. P. (1975). Logic and conversation. In Cole, P., editor, *Syntax and Semantics 3: Pragmatics*. Academic Press, New York.

Groenendijk, J. and Roelofsen, F. (2009). Inquisitive semantics and pragmatics. In Larrazabal, J. and Zubeldia,

L., editors, *Meaning, Content, and Argument: Proceedings of the ILCLI International Workshop on Semantics, Pragmatics, and Rhetoric*, pages 41–72. University of the Basque Country Publication Service.

Groenendijk, J. and Stokhof, M. (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD Dissertation, Institute for Logic, Language, and Computation (ILLC), University of Amsterdam.

Grosz, B. and Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Hamblin, C. L. (1973). Questions in montague english. *Foundations of Language: International Journal*, 10(1):41–53.

Harris, D. W. (2020). We talk to people, not contexts. *Philosophical Studies*, 177(9):2713–2733.

Hoek, D. (2018). Conversational exculpature. *The Philosophical Review*, 127(2):151–196.

Jankovic, M. (2014). Communication and shared information. *Philosophical Studies*, 169(3):489–508.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30:1–45.

Lederman, H. (2017). Uncommon knowledge. *Mind*, forthcoming.

Ludlow, P. (2014). *Living Words*. Oxford University Press.

MacFarlane, J. (2016). Vagueness as indecision. *Proceedings of the Aristotelian Society, Supplementary Volume*, 90(1):255–283.

MacFarlane, J. (2020). Indeterminacy as indecision, lecture 3: Indeterminacy as indecision. *The Journal of Philosophy*, 117(11/12):643–667.

Perrault, C. R. (1990). An application of default logic to speech act theory. In Cohen, P. R., Morgan, J., and Pollack, M. E., editors, *Intentions in Communication*, chapter 9, pages 161–185. MIT Press, Cambridge, MA.

Planer, R. J. and Sterelny, K. (2021). *From Signal to Symbol: The Evolution of Language*. MIT Press, Cambridge, MA.

Plunkett, D. and Sundell, T. (2013). Disagreement and the semantics of normative and evaluative terms. *Philosophers' Imprint*, 13(23).

Roberts, C. (2012a). Information structure: Afterward. *Semantics and Pragmatics*, 5(7):1–19.

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

Roberts, C. (2017). Linguistic convention and the architecture of interpretation. *Analytic Philosophy*, 58(4):418–439.

Roberts, C. (2018). Speech acts in discourse context. In Fogal, D., Harris, D., and Moss, M., editors, *New Work on Speech Acts*, Oxford. Oxford University Press.

Roberts, C. (2022). The architecture of interpretation: Dynamic pragmatics and pragmatic competence. *Unpublished manuscript of the 2022 Beth Lecture, University of Amsterdam*.

Schaffer, J. (2007). Knowing the answer. *Philosophy and Phenomenological Research*, 75(2):383–403.

Schoubye, A. and Stokke, A. (2016). What is said? *Noûs*, Online First.

Simons, M. (2001). Disjunction and alternativeness. *Linguistics and Philosophy*, 24:597–619.

Simons, M., Beaver, D., Roberts, C., and Tonhauser, J. (2017). The best question: Explaining the projection behavior of factives. *Discourse Processes*, 54(3):187–206.

Simons, M., Tonhauser, J., Beaver, D., and Roberts, C. (2010). What projects and why. *Semantics and Linguistic Theory*, 20:309–327.

Stalnaker, R. (1974). Pragmatic presuppositions. In Munitz, M. and Unger, P., editors, *Semantics and Philosophy*, pages 197–214. New York University Press, New York.

Stalnaker, R. (1978). Assertion. In Cole, P., editor, *Syntax and Semantics 9*, pages 315–332. Academic Press, New York.

Stalnaker, R. (1984). *Inquiry*. MIT Press, Cambridge, Mass.

Stalnaker, R. (2002). Common ground. *Linguistics and Philosophy*, 25:701–721.

Stalnaker, R. (2014). *Context*. Context and Content. Oxford University Press, Oxford.

Thomason, R. H. (1990). Accommodation, meaning, and implicature: Interdisciplinary foundations for pragmatics. In Philip. R. Cohen, J. M. and Pollack, M. E., editors, *Intentions in Communication*, chapter 16, pages 325–364. MIT Press, Cambridge, Mass.

Unnsteinsson, E. (2022). *Talking About: An Intentionalist Theory of Reference*. Oxford University Press, Oxford.

von Stechow, A. (1991). Focusing and backgrounding operators. In Abraham, W., editor, *Discourse Particles: Descriptive and Theoretical Investigations*, volume 12 of *Pragmatics & Beyond: New Series*, pages 37–84. John Benjamins, Amsterdam.

Yalcin, S. (2018). Belief as question-sensitive. *Philosophy and Phenomenological Research*, 97(1):23–47.