

# Relativism, Metasemantics, and the Future

Derek Ball  
University of St Andrews

**ABSTRACT:** Contemporary relativists often see their view as contributing to a semantic/post-semantic account of linguistic data about disagreement and retraction. I offer an independently motivated metasemantic account of the same data, that also treats a number of cases that are problematic for the relativist. The key idea is that the content of assertions and beliefs is determined in part by facts about other times, including times after the assertion is made or the belief is formed. On this *temporal externalist* view, speaker behaviours such as retraction of previous assertions play a role in making it the case that a past utterance has a given meaning.

Recent work by John MacFarlane (2005; 2009; 2011a; 2014), Mark Richard (2004; 2008), Max Kölbel (2002; 2004), and others has given new life to philosophical debate about relativism.<sup>1</sup> These new age relativists are motivated not by grandiose claims about the nature of truth itself, but by mundane linguistic facts about how speakers agree, disagree, and respond to new evidence. On MacFarlane's view, relativism is the claim that some expressions of natural language are *assessment-sensitive*, so that a single occurrence of a sentence containing such expressions can be true relative to some context of assessment and false relative to another; in particular, the relativist predicts that we should retract earlier assertions if the asserted sentence is false as assessed from our present context (even if it was true as assessed from the original context of assertion). Typically, this sort of relativist view is developed in a semantic framework in which sentences express propositions;<sup>2</sup> the relativist then claims that there is an important notion of propositional truth according to which propositions are true and false relative to some parameter that is supplied by a context of assessment. According to MacFarlane, this form of relativism is a part of a semantic/pragmatic theory of natural language, to be evaluated by empirical evidence:<sup>3</sup>

---

<sup>1</sup> Other work in this relativist renaissance includes Brogaard (2008, 2009); Egan et al. (2005); Egan (2007, 2010); Lasersohn (2005, 2008, 2009); Stephenson (2007); Weatherson (2009).

<sup>2</sup> MacFarlane (2014, ch. 3) points out that the relativist's view about when assertions should be defended or retracted need not be developed in a propositional framework. Though none of my arguments depend on this, for the sake of simplicity, I will focus on the propositional version of the view.

<sup>3</sup> MacFarlane (2003; 2008) offers another motivation for relativism, which depends on a controversial metaphysical view about the nature of time. For the purposes of this paper, I set

allowing accuracy to be assessment-sensitive has definite consequences for the predictions we make about when speakers will take themselves to be warranted in making assertions, when they will feel normative pressure to retract earlier assertions, and when they will take themselves to be in disagreement. Understood in this way, relativism about a particular domain of thought and talk is not a metaphysical thesis but a testable, empirical hypothesis — at least to the extent that any semantic theories are testable. (2011b, p. 444)

Although MacFarlane makes a good case that some areas of speaker behaviour are neatly explained by his relativism, there remain data that seem to fit relativism badly (see von Fintel and Gillies (2008), and section 1.2 below). Moreover, there are serious worries about whether relativistic accounts of assertion and belief are tenable (Wright, 2008; Zimmerman, 2007). Fortunately, there is an alternative account of the data. I will argue that the relativists' data (and von Fintel and Gillies's anti-relativist data) can be accounted for on certain externalist views of content. In particular, I argue that the relativist data can be explained if the content of assertions and beliefs is determined in part by facts about other times, including times after the assertion is made or the belief is formed. On this view, speaker behaviours such as retraction of previous assertions play a role in making it the case that a past utterance has a given meaning. *Pace* the scepticism exhibited by some critics of this temporal externalism (Brown, 2000), I will argue that standard externalist arguments of the sort developed famously by Burge and Putnam can be motivated in defence of my view. It will turn out that the MacFarlane cases share a structure with famous externalist cases, and that similar arguments can be applied.

The distinction between my view and the relativist's can be put as follows. David Kaplan (1989) distinguished *descriptive semantics*, which attempts to assign semantic values to expressions, from *metasemantics*, which attempts to give an account of how expressions get the semantic values they have. The typical relativist tries to explain a certain pattern of assertion and retraction in terms of a descriptive semantic claim (about the semantic values of sentences), which is used as part of a theory about when certain assertions are treated as accurate and inaccurate (which MacFarlane (2003, p. 329) calls "postsemantics"). In contrast, I claim that there is a metasemantic explanation of the same data. Since the metasemantic explanation is independently motivated, the relativist semantics and postsemantics is otiose.<sup>4</sup>

My claim, in short, is that MacFarlane-style relativism mistakes a metasemantic

---

this motivation aside.

<sup>4</sup>Glanzberg (2007) also defends the idea that the relativist data can be accounted for metasemantically, but does not consider temporal externalist views of the sort I develop here.

tic phenomenon for a semantic/postsemantic one. I defend this claim as follows. In section 1, I review the data that motivate relativism, as well as the data that cast doubt on the relativist's view. In section 2 I explain and defend temporal externalism. In particular, I try to show that minor variations on the cases and arguments used by Tyler Burge to defend the claim that meaning depends in part on the social can also be used to show that meaning also depends in part on facts about the future; indeed, Burge's cases are structurally quite similar to MacFarlane's. Section 3 shows that the temporal externalist has an account of the relativist data that also explains the von Fintel and Gillies data. Section 4 gives an account of assertion in a temporal externalist framework.<sup>5</sup> I conclude by suggesting that the temporal externalist account has wide significance: once we see the way in which metasemantic phenomena can be mistaken for semantic phenomena, numerous philosophical issues can be seen through this lens, including debates about conditionals, vagueness, the context-sensitivity of "knows", and the meanings of logical terms.

## 1 The Relativist Data

I will focus on two sorts of interactions between speakers that have been alleged to motivate relativism. The first is most associated with John MacFarlane. MacFarlane and others have applied this strategy to a number of areas of language; I will focus on the case of *epistemic modals* (for example, "The keys might be on the table," where this is meant to indicate not that it is (e.g.) physically or metaphysically possible that the keys are on the table, but something along the lines of that it is compatible with the speaker's knowledge that they are). The second sort of case for relativism is due to Mark Richard. I will follow Richard in focusing on the case of *gradable adjectives*, like "rich" and "tall".

### 1.1 MacFarlane's Data

John MacFarlane (2011a) has argued that relativism is uniquely well-positioned to handle cases involving *retraction*, such as (1):

- (1) **Bob:** Donna might be delivering Meals on Wheels.  
**Mike:** No, you're wrong, she can't be. I just saw her at the diner.  
**Bob:** Oh, okay. What I said before was false.

MacFarlane reasons roughly as follows. Suppose that for all Bob knows at the time of his first utterance, Donna is delivering Meals on Wheels. Then his ut-

---

<sup>5</sup>Jackman (1999, 2005) offers the only detailed development of temporal externalism in the current literature. But Jackman focuses on motivating temporal externalism by appealing to thought experiments; my discussion goes well beyond Jackman in developing the metaphysics of meaning, pragmatics, and applications of temporal externalism.

terance seems appropriate. But it also seems appropriate for Mike to say that Bob spoke falsely, and for Bob to retract his assertion on the basis of Mike's new information. So what seems to be needed is an object that is appropriate for Bob to assert, for Mike to deny, and for Bob to retract.

On one traditional view, the objects of assertion are sets of possible worlds, or, equivalently, functions from worlds to truth values (e.g., Stalnaker (1984)). But no such object seems to meet the desiderata. Call the time of Bob's first utterance  $t_1$ , the time of Mike's utterance  $t_2$ , and the time of Bob's second utterance  $t_3$ . Consider the (possible-worlds) proposition that it is compatible with what Bob knows at  $t_1$  that Donna is delivering Meals on Wheels – that is, the set of worlds  $w$  such that it is compatible with what Bob knows at  $t_1$  in  $w$  that Donna is delivering Meals on Wheels. This proposition is true, and Bob is in a good position to know that it is true; so it seems like an appropriate thing for Bob to assert. But it doesn't seem appropriate for Mike, and later Bob, to deny. In particular, it isn't appropriate for Mike to say "You're wrong", and it isn't appropriate for Bob to say, "What I said before was false." After all, if Bob asserted this proposition, then he spoke truly, what he said was true, and both he and Mike are in a position to know this. So the view that Bob asserted this proposition makes no sense of Mike and Bob's later interaction. But no other (possible worlds) proposition seems much better. For example, consider the proposition that it is compatible with what Mike and Bob know at  $t_2$  that Donna is delivering Meals on Wheels – that is, the set of worlds  $w$  such that it is compatible with what Mike and Bob know at  $t_2$  in  $w$  that Donna is delivering Meals on Wheels. This is appropriate for Mike to deny at  $t_2$ , since it is false and he is in a position to know that it is false. But it seems inappropriate for Bob to assert at  $t_1$ . After all, Bob might have no idea what Mike knows. And the objection can be strengthened by reflecting on the fact that it seems appropriate even for eavesdroppers to react as Mike does; it seems unlikely that that Bob would intend the knowledge of eavesdroppers to be relevant, and irresponsible if he did (given that he might have no idea who is listening in.)<sup>6</sup>

Similar objections seem to plague any possible-worlds proposition. MacFarlane concludes that the relevant object of assertion and denial is not a function from possible worlds to truth values, but a function from possible-worlds *and information states* (i.e., sets of worlds) to truth values. One may assert such an object only if it is true relative to one's own information state – typically, the worlds compatible with what one knows – and one may evaluate the assertions of others relative to one's own information state. This view makes sense of the dialogue between Mike and Bob. Bob asserts the proposition that is true relative to an

---

<sup>6</sup>Eavesdroppers are the primary motivation for relativist treatments of epistemic modals in Egan et al. (2005); Egan (2007). I return to the issue of eavesdroppers in section 3 below.

information state and a world just in case it is compatible with that information state in that world that Donna is delivering Meals on Wheels. This proposition is appropriate for him to assert: it is true relative to his new information state and his world at the time he asserts it, and he knows it to be so. But since Mike knows that Donna is not delivering Meals on Wheels, it is false relative to Mike's information state. So it is appropriate for Mike to reject. And when Mike rejects it, Bob learns that Donna is not delivering Meals on Wheels. Once he learns this, the proposition is false relative to his information state; so it is appropriate for him to retract his assertion.

### 1.2 Von Fintel and Gillies's Data

The story so far is an attractive one. But there is other data that the relativist cannot handle so smoothly. Von Fintel and Gillies (2008) point out that the following dialogue is equally natural:

- (2) **Bob:** Donna might be delivering Meals on Wheels.  
**Mike:** No, you're wrong, she can't be. I just saw her at the diner.  
**Bob:** I didn't say that she *was* delivering Meals on Wheels, only that she *might be*, and I was right: she might have been. Sheesh.

Bob's reply seems appropriate; it makes sense for him to stick to his guns here. But the relativist seems to predict that Bob's response is inappropriate. Given what Bob knows at the time of his reply, the proposition that Donna might be delivering Meals on Wheels is false relative to his information state. The pattern of reasoning that the relativist used to explain the first example should lead us to expect Bob to retract the claim that Donna might be delivering Meals on Wheels. The fact that Bob's failure to retract is completely natural and apparently correct suggests that something has gone wrong in the relativist's account.

In fact, the situation is not quite so simple: as MacFarlane (2014) points out, the relativist can argue that a sentence like "Donna might be delivering Meals on Wheels" can be used to assert that for all the speaker knows, Donna might be delivering meals on wheels, by appealing to the general phenomenon that we often speak loosely and omit such qualifications when we can expect an audience to pick up on our intentions (2014, p. 259). For example, MacFarlane suggests that von Fintel and Gillies example is much like the following:

- (3) **Kai:** It's 6:15.  
**Thony:** No, you're wrong, I just checked my watch and it's precisely 6:16.  
**Kai:** All I meant to assert was that it was *around* 6:15, and I was right: it is. Sheesh.

If MacFarlane is right, although relativism does not predict the existence of "stick to your guns" cases, they are unproblematic because they are to be expected on

independent grounds.

*Contextualist* opponents of relativism typically claim that the proposition expressed by a sentence involving an epistemic modal depends on context. The relativist's original case refuted the claim that an utterance of 'might  $\phi$ ' is true just in case  $\phi$  is compatible with what the utterer knows at the time of utterance. But more sophisticated stories are possible that can go a long way toward duplicating the relativist's results. The standard view of modals like "might" has it that they are quantifiers over worlds (the *locus classicus* of the standard sophisticated variant of this view in natural language semantics is Kratzer (1977)); thus 'might  $\phi$ ' says that there is some world in which  $\phi$  is true. But we typically are only interested in some limited subset of the worlds. So modal quantification is typically restricted. In the case of *epistemic* modals, the relevant restriction is to worlds compatible with someone's knowledge. But whose? The question is especially pressing in light of the existence of examples of examples such as (4), which seem to show that the speaker's knowledge need not be what is at issue (see Egan et al. (2005)):

- (4) **Situation:** Brian is worried that John is stalking him. As the number 12 bus approaches, Brian rushes to hide behind a bush.

**Andy:** Why did he hide?

**John:** I might be on that bus.

John knows well that he is not on the bus; so John's utterance only makes sense if we construe it as saying something like: it is compatible with what Brian knows that I am on the bus.

So epistemic modals can quantify over worlds compatible with the knowledge of individuals other than the speaker. And it might seem that this fact gives the contextualist all the resources she needs to explain the relativist's data. For the contextualist can claim that in the relativist's original case, the modal quantifies over worlds compatible with what is known either to Mike or to Bob; while in von Fintel and Gillies's "stick to your guns" case, the modal quantifies only over Bob's knowledge (at the time of utterance). (See Dowell (2011) for a defence of this sort of view.) This would explain why it is appropriate for Mike to retract in the first dialogue: his utterance was false, because there are no worlds compatible with what Mike knows in which Donna is delivering Meals on Wheels. And it would explain why it is not appropriate for Bob to retract in the second dialogue: his utterance was true, because there are worlds compatible with what he knows in which Donna is delivering Meals on Wheels.<sup>7</sup>

---

<sup>7</sup>Of course, a story would also have to be told about why it is appropriate for Mike to reject Bob's assertion. (Perhaps it can be argued that Mike is targeting not the claim that Donna might be delivering Meals on Wheels, but rather the *prejacent* – i.e., the claim that Donna *is* delivering Meals on Wheels (von Fintel and Gillies, 2008).)

Attractive though they may seem, both MacFarlane's relativism supplemented with claims about loose speaking and the flexible contextualist strategy are subject to serious objections. A first point against flexible contextualism is that the account does not evade the objections raised against earlier contextualist accounts. After all, on this view, the truth of Bob's first assertion in (1) depends (in some cases) on what Mike knows. But Bob's assertion seems appropriate even if Bob has no idea what Mike knows. (Recall again that Mike might be an eavesdropper who inserts himself into the conversation.)

But there is a further, crucial datum that is more difficult for either of these views to accommodate – as long as we are assuming a traditional metasemantics, where facts about the time of utterance are sufficient to fix the meanings of our words. MacFarlane suggests that one can stick to one's guns if an assessment-invariant proposition "is all [one] intended to assert in the first place" (2014, p. 259). Similarly, on the contextualist's view, a traditional (e.g., Gricean) metasemantic story would entail that it is the speaker's intentions at the time of utterance that fix the content of the modal; closely related views would appeal to a speaker's dispositions at the time of utterance. Both of these views would have the following consequence:

**Fixed at Time of Utterance (FTU)** Whether it is appropriate to retract or stick to one's guns in the face of correction is fixed at the time of utterance.

In particular, the proponent of a MacFarlane-style supplemented relativism or a flexible contextualism who endorses such a standard metasemantic story should hold that whether Bob's utterance makes an assertion of one of the (possible worlds) propositions that for all Bob knows Donna is delivering Meals on Wheels, or that for all Mike and Bob together know Donna is delivering Meals on wheels, or the proposition that is true relative to an epistemic state just in case it is compatible with that state that Donna is delivering Meals on Wheels, depends on what Bob intends (and perhaps also what intentions he can reasonably expect his audience to discern). In either case, it is fixed at the time of utterance whether he should retract or stick to his guns when faced with Mike's reply.

There are many cases in which (FTU) is plausible; for example, if Bob has a clear and unwavering intention that only his own knowledge is relevant to the truth of his utterance, and Mike is in a position to recognise this intention. In this sort of case, Mike's attempt to correct Bob seems like a mistake; Bob would be justified in sticking to his guns, and mistaken if he were to retract. Alternatively, Bob might have a firm intention that Mike's knowledge is also relevant. In this case, Mike's correction seems appropriate, and Bob must retract rather than sticking to his guns. But not every case is of this kind. In many – perhaps most – cases, our intentions are much less clear. (For example, perhaps Bob intends to make an assertion he knows to be true, but also an assertion whose truth depends

on what Mike knows – and these intentions may not be jointly satisfiable. I discuss a variety of further cases in section 2.1, below.) In many cases, we will not have resolved in advance how to respond to correction; we might have no particular intentions – even implicit ones – about whose knowledge might be relevant to the truth of a particular claim, or we might have conflicting intentions.

Does appealing to dispositions help? It may in some cases; some speakers may have fixed dispositions to retract or to stick to their guns. But many typical speakers will have much more complex dispositions: Bob might be disposed to retract if Mike objects politely, but to stick to his guns if Mike is rude or confrontational. Or Bob might be disposed to retract if he is in a good mood, but to stick to his guns if he is grumpy. (Plausibly, most of us have dispositions of this sort in very many conversational circumstances.) But there does not seem to be anything that could make it the case that one or the other of these dispositions is decisive. So it is hard to see how dispositions could do the relevant metasemantic work in such cases.

Suppose that Bob has such complex, indeterminate, or conflicting intentions and dispositions. The most natural construal of this case is one on which *either* response is permissible. Nothing at the time of his original utterance will have determined a uniquely correct response in advance. The possibility that Mike might object could be very far from his thoughts. Faced with the objection, he might decide on the spur of the moment to be concessive and to retract his original assertion. If he does so, both Mike's objection and Bob's retraction seem correct.<sup>8</sup> On the other hand, he might choose instead to stick to his guns. If he does so, this seems correct, and Mike's objection seems like a mistake – a mistake that could be registered by the participants in the conversation:

- (5) **Bob:** Donna might be delivering Meals on Wheels.  
**Mike:** No, you're wrong, she can't be. I just saw her at the diner.  
**Bob:** I didn't say that she *was* delivering Meals on Wheels, only that she *might be*, and I was right: she might have been. Sheesh.  
**Mike:** Oh, I see. Then I was wrong; what you said before was true.

In this sort of case, then, we could either stick to his guns or retract; either response would make sense, and neither would seem incorrect. We may *choose* how to respond. Our responses are not determined in advance. Thus (FTU) cannot be correct.

---

<sup>8</sup>Of course, someone might grant that such a response is in some sense appropriate – they are in keeping with at least some conversational norms – while denying that it is true. But this response is not available to the relativist, or to the sophisticated contextualist who accepts the relativist data: the same response could have been made to MacFarlane's original examples.



The data, then, seems to pattern as follows:<sup>9</sup>

**The Data Pattern** There are circumstances under which:

1. A speaker who does not know whether an utterance of  $\phi$  by her would express a truth may reasonably make an assertion using a sentence of the form ‘It might be that  $\phi$ ’.
2. An interlocutor who does know whether an utterance of  $\phi$  by the first speaker would have expressed a truth may respond by rejecting the first speaker’s assertion (e.g., by saying “No, you are wrong”, or “No, you spoke falsely”, or ‘No, it cannot be that  $\phi$ ’.)
3. The original speaker may make either of two sorts of response: she may either retract her original assertion, or reiterate her original assertion. If she retracts, then this retraction seems correct (and her original assertion seems incorrect). If she reiterates, then her reiteration seems correct (and her original assertion also seems correct).

Although this way of describing the data is quite natural and suggestive, characterising the phenomenon precisely is a subtle matter, and the proponent of a traditional metasemantics might object to the way I have described things so far. In subsequent sections, I develop a number of other cases that exhibit a similar pattern. I will argue that the resources needed to explain these cases can explain the data about epistemic modals as well, and that a unified explanation of this kind is highly attractive. But first, I turn to another sort of contextualist account that purports to explain the data presented so far.

### 1.3 Richard’s Data

Richard argues that it is possible for two people to disagree – and for their disagreement to be substantive rather than merely verbal – about (for example) whether someone is rich, despite agreeing about that person’s level of wealth and about the interests and purposes relevant to the conversation (including, for example, a comparison class relative to which degree of wealth is being evaluated.) Thus Richard claims that two people might argue about whether Didi is rich, despite agreeing that Didi makes exactly \$250,000 per year, and agreeing that what is at issue is whether she is rich for a New Yorker; one might point out

---

<sup>9</sup>Von Fintel and Gillies’s own view of the cases involves the idea that utterances take place not in a single determinate content, but in many contexts, so that many propositions are “put into play” by a typical utterance. The sort of formal apparatus that they employ could be used to model a temporal externalist view. But the norms of assertion and confirmation or denial that they propose (2011, p. 120-1) are unacceptable for essentially the same reasons as the relativist and contextualist views discussed: they presuppose that what propositions are put into play is fixed at the time of utterance, and so cannot accommodate this data pattern.

that \$250,000 enables one to buy many luxuries, while the other might argue that compared to the wealthiest New Yorkers, \$250,000 is really not so much money at all. (Richard claims further that they could disagree even though both the utterer of “Didi is rich” and the utterer of “Didi is not rich” could each correctly regard her own utterance as true. I return to this claim below.)

Richard suggests that part of what is at issue in such a dispute is what David Lewis (1979) called *accommodation*. “Rich” is a gradable adjective. Gradable adjectives are characteristically context sensitive: just as what counts as “tall” for professional basketball players is different than what counts as “tall” for kindergarteners, what counts as “rich” for professional philosophers is different than what counts as “rich” for investment bankers. What determines the contextually relevant standard of height, or of wealth? Lewis suggested that the standards are determined in part by what we say, and in particular that the context tends to be adjusted to make our utterances true. Thus for example, if I say, “My five year old cousin is very tall,” a context tends to be created in which comparison to other five-year-olds is salient, and 3’ 9” counts as “tall”. If, on the other hand, I say, “Former NBA star Allen Iverson is not tall,” a context tends to be created in which comparison to other professional basketball players is salient, and 6’ does not count as “tall”.

Of course, the view is not that whatever anyone says goes. For example, Lewis claims that presuppositions (such as the presupposition of “The king of France is bald” that France has one king) are accommodated if, but only if, “conversational partners tacitly acquiesce – if no one says ‘But France has *three* kings!’ ” (1979, p. 234). Conversations are typically cooperative, but they need not be; one typically accommodates, but one can refuse. We need a theory of what happens when one party refuses to accommodate. On Richard’s view, it is here that relativism plays a role. Consider again the debate about whether Didi is rich:

(6) **Dilip:** Didi is rich. She took me out to a really fancy restaurant.

**Torfinn:** Didi is not rich. She can’t afford a penthouse on Park Avenue.

The participants in such a debate seem genuinely to disagree;<sup>10</sup> moreover, they seem to disagree about whether Didi is rich. (In particular, although they may in some sense be disagreeing about the meaning of the word “rich” in this context, that is not all they are disagreeing about; after all, the reasons that they cite (e.g., that Didi can afford a fancy restaurant) bear on Didi’s lifestyle, not (or not directly) on the semantics of English).

---

<sup>10</sup>Cappelen and Hawthorne (2009, pp. 60-1) distinguish between disagreement as an activity (potentially involving, e.g., claims of “You’re wrong”, raised voices, etc.) and disagreement as a state. In the stative sense, we can attribute disagreement to people who do not know of each other’s existence, who live at different times and places and never meet, etc. It is the stative sense that is at issue here, in the Burge arguments discussed below, and throughout this paper.

The puzzle can be put in the following way. The sort of disagreement that underlies the debate only makes sense if there is a proposition over which the parties disagree; that is, a proposition which one party accepts and the other denies. What proposition could this be? It cannot be the proposition that Didi is rich enough to eat at fancy restaurants; Torfinn would not disagree with this proposition. Nor could it be a proposition that builds in a specific standard of wealth – say, that Didi is richer than people who make \$200,000 per year. There is no proposition of this sort that Dilip would affirm and Torfinn would deny. And in general, if propositional truth is construed as being relative only to worlds, there seems to be no good candidate.

Richard argues that we come to a solution if we embrace relativism. For then we can hold that Dilip and Torfinn disagree over a common proposition: the proposition that Didi is rich (or perhaps the proposition that Didi is rich for a New Yorker). This proposition is true and false only relative to standards of wealth (or perhaps something that determines such standards, such as judges). Thus Dilip accepts this proposition relative to his standards (according to which someone is rich for a New Yorker if she makes more than \$200,000 per year), and Torfinn rejects this proposition relative to his standards (according to which someone is rich for a New Yorker only if she makes more than \$600,000 per year). On this view, the debate is about the truth of the proposition that Didi is rich. The debate will be resolved if one party can convince the other to change his or her standard.

Richard calls such debate “negotiation”. On Lewis’s view, accommodation is a matter of changing the context. Such changes are typically temporary: they last only as long as the conversation. But on Richard’s view, accommodation and negotiation play a deeper role. Richard claims that what is at issue in cases of accommodation and negotiation is “how to shape the boundaries of a concept” (2008, p. 118). That is, the goal is not to change the context, but to change the standing meaning.<sup>11</sup> (One might say: Lewis describes how to keep track of changes in the score in a language game; Richard suggests that sometimes we are trying to change the rules of the game.) The participants in this sort of debate are suggesting systems of classification that would be useful. One may make an assertion using a sentence  $\phi$  in such a case not because our current linguistic practices make it the case that  $\phi$  is correctly applied the situation at hand, but rather because one is suggesting that we ought to shape our linguistic practices so as to make it the case that  $\phi$  is correctly applied to such cases.

---

<sup>11</sup>Note that I am not insisting that Richard is right about “rich”, though I will insist that there are relevantly similar cases (such as the case of “mass”, discussed in the next paragraph) where it is clear that the upshot of negotiation is change in standing meaning. One advantage of the account I supply below is that it can be applied both to standing meaning and to contextually supplied parameters. See section 3 below for further discussion.

Richard's view, then, has two components:

- Accommodation and negotiation are a matter of trying to find a common standard at which to evaluate relatively true propositions.
- Accommodation and negotiation are a matter of shaping our concepts.

Given the first component, we can reasonably take cases such as the “rich” case described above as involving non-metalinguistic disagreement. And surely there are some cases to which the second component applies, whether or not the “rich” case is among them. It is very plausible that our words and concepts are open textured (in the sense of Waismann (1945)): there are cases for which our antecedent usage and other meaning-relevant factors (e.g. causal history) do not determine whether a given word or concept applies or fails to apply. In these cases, there is an element of semantic decision: we could choose to extend our practice so that the word applies in the case, or extend the practice so that it does not apply. (I will return to the issue of whether this sort of decision make for a change of meaning.) Consider a case discussed by Hartry Field:

Consider the use of the term “heavier than” by pre-Newtonians. Did it stand for the relation of having greater mass than, or for the relation of having greater weight than? In pre-Newtonian physics there was no distinction between the weight of an object and its mass; and since the term “heavier than” was applied almost exclusively in the context of objects at the surface of the earth where there is a near-perfect correlation between mass and weight, there is little in the pre-Newtonian use of the term that could have settled the matter. (2001a, p. 278)

If Field is right, nothing in pre-Newtonian usage settled the matter of whether (say) a feather on Earth is heavier than an Olympic barbell in space. The issue had never arisen; neither way of proceeding would contradict previous usage. (We can suppose further that no one had intentions or dispositions to apply the term in one or another of these ways, or that there are inconsistent or confused intentions and dispositions across the community.)

Field (2001b) originally proposed the example as a case where it is indeterminate which relation “heavier than” picks out; he proposes an account of *partial designation* on which we can say that it partially picks out *having greater weight than* and partially picks out *having greater mass than*. Given our stipulations about the case, some sort of partiality or indeterminacy seems to be mandated if we assume that facts about usage (intentions, dispositions, etc.) at or before the time of utterance are what determines meaning. But it is not obvious that this

account can make sense of the fact that one can imagine Richardian negotiation playing a role in these semantic decisions. For example, one might imagine various Newtonians debating the question of whether the feather or the barbell is heavier, giving reasons and arguments to support one side or the other (perhaps the proponents of *greater weight* would point out that the two objects could be lifted with equal ease, while the proponents of *greater mass* would point out that the barbell would exert greater force (at the same rate of acceleration)). And one can imagine one side or the other backing down in the face of stronger arguments, retracting her assertions, seemingly correctly saying things like “I was wrong”, and so forth. But if it is indeterminate whether “heavier than” picks out *having greater weight than* or *having greater mass than*, neither party to the debate is (determinately) correct; and if designation in this case is partial, both sides to such a debate would be partially wrong. So on this sort of view, the debate and retraction behaviour would be confused. Since Richardian negotiation makes sense, the partial/indeterminate reference view should be rejected.

Richard’s view also fails to explain the case. Negotiation, for the relativist, shapes the context at which we apply our concepts. Thus, to the extent that Richardian negotiation is a matter of changing the context at which relatively true propositions are evaluated, accommodation and negotiation do not change our concepts at all. (So despite his rhetoric about concept change, accommodation on Richard’s view is little different than accommodation on Lewis’s view.) There certainly may be cases in which context change is all that is at issue. But this does not seem to be what is at work in the open texture cases. And it does not seem to be what is at work in Richard’s “rich” case, granting Richard’s own view that the case is about shaping the concept of being rich for a New Yorker.

In short, the situation is this. Richard’s account of accommodation and negotiation turns on postulating non-standard semantic values (i.e., propositions that are only relatively true) for sentences containing expressions like “rich”. But this is surprising. Recall Kaplan’s distinction between descriptive semantics and metasemantics. Negotiation, as Richard describes it, is a matter of shaping our concepts: that is, determining the semantic value of a term or concept. It thus seems like a *metasemantic* phenomenon. So it would be surprising if a strictly *semantic* account of the sort Richard gives is adequate.

In the next section, I begin to lay out the metasemantic framework that will explain Richardian negotiation. But before doing so, we should note that Richard’s cases fit the Data Pattern described with respect to the case of epistemic modals above. But Richard’s cases are different from the case of epistemic modals because they allow for *negotiation* by each participant in the conversation.<sup>12</sup> First, once the original speaker declines to retract, the objector is faced

<sup>12</sup>I return to the case of epistemic modals (where further debate seems strange and inappro-

with a similar choice: she may retract her rejection of the original utterance, or she may reiterate it. If she reiterates, then the original speaker must again choose. Second, at each stage both the original speaker and the objector may offer reasons for their view (as well as responses to reasons offered by the other speaker). A speaker who chooses to retract may do so because of the reasons offered by the other party. (Thus the choice to retract is not merely arbitrary or even prudential, but can be rational in an epistemic sense.) In sum:

**The Data Pattern** There are circumstances under which:

1. A person who associates being rich for a New Yorker with a certain standard of wealth may reasonably make an assertion using a sentence of the form “Didi is rich”.
2. A person who associates being rich for a New Yorker with a different standard of wealth may respond by rejecting this assertion (e.g., by saying “No, you are wrong”, or “No, you spoke falsely”, or “No, Didi is not rich”.)
3. The original speaker may make either of two sorts of response: she may either retract her original assertion, or reiterate her original assertion. If she retracts, then this retraction seems correct (and her original assertion seems incorrect). If she reiterates, then (modulo (5) below) her reiteration seems reasonable (and her original assertion continues to seem reasonable).
4. If the original speaker reiterates, the objector may make either of two sorts of response: she may either retract her rejection, or reiterate her rejection. If she retracts, then this retraction seems correct (and her original objection seems incorrect, and the other speaker’s original assertion seems correct). If she reiterates, then (modulo (5) below) her reiteration seems reasonable (and her original rejection continues to seem reasonable).
5. (3) and (4) may be repeated indefinitely. At each stage, a speaker may offer reasons for her view, and must respond to reasons offered by her interlocutor.

In the next section, I turn to a metasemantic account that promises to explain this Data Pattern.

---

priate) in section 3 below.

## 2 Temporal Externalism

In the *Nicomachean Ethics* I.9-10, Aristotle discusses the view that there is no fact of the matter as to whether a man is leading a happy life until after he is dead: “there is required not only complete virtue but also a complete life, since many changes occur in life, and all manner of chances, and the most prosperous may fall into great misfortunes in old age, as is told of Priam in the Trojan Cycle; and one who has experienced such chances no one calls happy” (NE I.9). In particular, “we must see the end and only then call a man blessed, not as being blessed but as having been so before” (NE I.10). Thus, on this view, whether a person is happy at age 30 depends in part on what happens to her at age 60.

Aristotle attributes this view to Solon. Say that a property  $p$  is *Solonic* if and only if whether an object has  $p$  at a time  $t$  depends in part on what happens after  $t$ . On the view Aristotle discusses, the property of being happy (or blessed) is Solonic. Other examples of Solonic properties include the property of being the next President of the United States, the property of being the beginning of a beautiful friendship, the property of being the day Jack quit smoking, the property of knowing what will happen, the property of being about to die, and the property (often attributed to Michael Jordan) of being the G.O.A.T. (the greatest [basketball player] of all time).

The temporal externalist holds that properties like *meaning that water is wet* and *believing that arthritis is a disease* are Solonic. How does this connect to more familiar formulations of externalism about meaning and content? Externalism is often characterised in terms of supervenience: a set of properties A supervenes on another set of properties B just in case no two objects can differ with respect to their A properties unless they differ with respect to their B properties; the internalist holds, while the externalist denies, that intentional properties like *believing that arthritis is a disease* supervene on purely intrinsic properties of a thinker. Equivalently, if the A properties supervene on the B properties then we can say that the B properties are the supervenience base of the A properties; in this terminology, the internalist holds, while the externalist denies, that there is a supervenience base of intentional properties that consists of purely intrinsic properties of a speaker or thinker.

On certain assumptions, temporal externalism can be spelled out in this way: temporal externalists hold that the intentional properties an object has at a time supervene not only on properties that the object has at or before that time, but also on the properties it has at later times. There are two reasons for caution, which I will mention only to set them aside. First, if determinism is true, then the course of the future is necessitated by the facts about the present (including facts about natural laws). But then every world where all the facts about the present are the same is a world in which all the facts about the future are the same. So

the facts about the future supervene on the facts about the present. And so the supervenience of content on the future would reduce to supervenience on the present. In this case, the temporal externalist can either try to formulate her view (and her opponents') in terms of a more specific supervenience base (so that, for example, the temporal internalist would affirm that intentional properties supervene not just on all matters of fact in the world up to the time of utterance, but on some subset of these, while the temporal externalist would maintain that some future facts must be included in the supervenience base), or to formulate her view in terms of some other sort of metaphysical dependence. (For some options, see Schaffer (2009); Rosen (2010).) Second, some presentists might deny that there are facts about later times. I will assume for the purposes of this paper that presentists have some way of making sense of this manner of speaking.<sup>13</sup>

Why think that temporal externalism is true? A preliminary reason is that it is entailed by popular views about the possibility of introducing directly referential terms by using a description to fix their reference. For example, Kaplan suggested that one could now assert things of the first child born in the 22nd century using his "Dthat" operator (which transforms a description into a Millian singular term that refers to the thing described); alternatively, we could arguably introduce the name "Newman-1" by stipulating that "Newman-1" is to name the first person born in the 22nd century. Kaplan wrote, "It is now clear that I can assert *of* the first child to be born in the twenty-[second] century that *he* will be bald, simply by assertively uttering, 'Dthat ['the first child to be born in the twenty-[second] century'] will be bald'" (1996, p. 303). Given that two utterances have the same meaning only if they refer to the same entities, then if Kaplan is right, the meaning of his utterance depends on what happens in the first moment of the 22nd century. So temporal externalism is true.

One would not wish to rest too much on these unusual and controversial cases. Fortunately, I think that a stronger case can be made. In order to make this case, I will present Tyler Burge's arguments for the claim that meaning depends on the social in some detail. These arguments will play two roles in the discussion to follow. First, simple variations on Burge's cases can, in my view, be used to generate an argument for temporal externalism. However, I have encountered some readers who do not share my judgments about these cases, and the argument for temporal externalism in Burge-style scenarios is independent of the argument for temporal externalism in the MacFarlane, von Fintel and Gillies, and Richard scenarios that we have already discussed. The second

---

<sup>13</sup>In order to avoid misunderstanding, let me be clear that although I am speaking of meaning as depending on "the future", temporal externalism does not depend on the existence of A-properties such as the property of being future. It can be specified equally well as the view that the meaning of an utterance at *t* depends in part on what happens at times later than *t*.



role that the discussion of Burge will play is to introduce a style of argument that can be used to defend for temporal externalism in the relativists' cases. I go on to make this argument in section 3.

### 2.1 Burge on "Arthritis"

Burge (1979; 1988) famously argues against internalism on the basis of thought experiments in which a character over- or under-extends a word. Burge's most famous case has us imagine a person who, in addition to a number of normal layperson beliefs about arthritis, has a belief that she would express by saying, "Arthritis can occur in the thigh". When he feels a pain in his thigh, he becomes suspicious about its cause, and reports to his doctor, "I might have arthritis in my thigh." The doctor corrects him and he accepts the correction.

Burge argues that the best explanation of the naturalness and seeming rationality of Arthritis-man's behaviour is that Arthritis-man means *arthritis* with his uses of "arthritis", and similarly possesses the same concept of arthritis as the expert, and so believes that arthritis can occur in the thigh. He argues that no other view can explain Arthritis-man's behaviour as well. In particular, he argues against the view that Arthritis-man uses the word "arthritis" with a different meaning than the doctor, and has a corresponding idiosyncratic belief (for example, the belief that *tharthritis* can occur in the thigh, where "tharthritis" has in its extension both diseases that occur in joints and diseases that occur in thighs), combined with the belief that the sentence "Arthritis can occur in the thigh" is typically used to express the claim that tharthritis can occur in the thigh. He has two main arguments against this view:

**Disagreement** We naturally attribute to Arthritis-man attitudes such as agreement and disagreement with speakers with more normal attitudes, and even with experts.<sup>14</sup> Moreover, we naturally regard him as sharing beliefs with such speakers. For example, he could agree with a doctor that arthritis is a disease, and in such a case he and the doctor would both believe that arthritis is a disease. This could be true even if the doctor in question is a monolingual speaker of Swahili, with whom Arthritis-man shares no relevant metalinguistic beliefs. Those who claim that Arthritis-man uses "arthritis" with an idiosyncratic meaning must deny that Arthritis-man agrees and that he disagrees with experts, and those who claim that he has an idiosyncratic belief must deny that Arthritis-man shares beliefs with the doctor. But this denial is implausible.

**Response to Correction** Arthritis-man accepts the doctor's correction. Later, he

---

<sup>14</sup>For a defence of the idea that attributions of agreement and disagreement are a good test of shared meaning and content, see Cappelen and Hawthorne 2009, ch. 2.

might reflect, “I used to think that I had arthritis in my thigh. What a silly mistake!” The idiosyncratic meaning/metalinguistic belief strategy predicts that this utterance is false: he used to believe that he had *tharthritis* in his thigh, and that belief was true. It is unclear why Arthritis-man would make such an error. Moreover, the idiosyncratic/metalinguistic strategy makes it unclear why Arthritis-man would accept his doctor’s correction in the first place. It seems to predict that Arthritis-man should respond to the doctor: “I know what disease is in my thigh; you’re just quibbling about the meanings of words.”

Of course these arguments are not decisive proofs. It is open to the internalist to reject the data (that Arthritis-man agrees and disagrees with others; that he is rational in responding to correction in the way he does), or to propose some alternative explanation of the data (perhaps in terms of similarity of content). But in the absence of well worked-out versions of such explanations, (and given the seeming difficulty of such a project), Burge proposes that the externalist view is the best explanation of the data. And this style of argument will be generalizable to any situation with similar patterns of agreement and disagreement and response to correction.

### **2.1.1 Do Intentions and Dispositions at the Time of Utterance Determine Content?**

One natural account of the Burge phenomena is that one’s intentions or dispositions (at the time of utterance) to accept the correction of relevant experts play a decisive role in making it the case that one means what the experts mean.<sup>15</sup> For example, it is important that Arthritis-man is disposed to accept the correction of his doctor; if he is disposed to stick to his idiosyncratic usage, then (other things being equal) the Burge arguments seem irrelevant. It is tempting to suppose that intentions and dispositions do all the work in making it the case that we mean what we do. But often, speakers’ intentions and dispositions are complex and conflicted. Arthritis-man intends to mean what the experts mean, but he also intends to pick out the disease in his thigh, and these intentions cannot both be satisfied. What makes it the case that the deferential intention is the decisive one?

In the most natural way of developing the case, it is plausible that the intention to use “arthritis” to pick out diseases of the thigh depends on the more fundamental intention to use “arthritis” as the experts do (and the belief that the

---

<sup>15</sup>There are, of course, familiar Wittgensteinian reasons to doubt that intentions and dispositions are sufficient to play this meaning-determining role. And the role of future decision and ongoing behaviours or practices in determining what one meant all along are important themes in Wittgenstein’s writing on rule following: see for example Wittgenstein 1953, §186, Wittgenstein 1967, pp. 77, 122, Wittgenstein 1976, pp. 30-31, 124, 237-8.

experts apply “arthritis” to diseases of the thigh). This is reason to think that the deferential intention is more important. But not every case is like this. It is possible to imagine Arthritis-man with independent and equally strong intentions, and no lasting dispositions to favour one over the other. In that case, it is *prima facie* plausible to maintain that Arthritis-man’s use of “arthritis” is indeterminate in meaning. Jessica Brown (2000) has developed a theory that has this result. Brown’s theory works as follows. Suppose speakers at time  $t_o$  use a word  $w$ , and consider some object  $o$  which may or may not be in the extension of  $w$  (as used at  $t_o$ ). Such speakers will have intentions about their use of  $w$ . For example, they might have *de re* intentions of some objects  $m$  and  $n$  that they are in the extension of  $w$ , or *de dicto* intentions that all  $F$ s are in the extension of  $w$ . If a speaker intends that  $o$  be in the extension of  $w$ , or if  $o$  is  $F$  and the speaker intends all  $F$ s to be in the extension of  $w$ , say that the speaker’s intentions *determine* that  $o$  is in the extension of  $w$ . And if the speaker’s intentions determine both that  $o$  is in  $w$  and that  $o$  is not in  $w$  (for example, suppose that they intend that  $o$  not be in the extension of  $w$ , but that all  $F$ s be in the extension of  $w$ , failing to realise that  $o$  is  $F$ ), say that the speaker’s intentions generate a conflict. Brown poses a dilemma: either speakers have intentions (at  $t_o$ ) as to whether  $o$  is in the extension of  $w$  as they use it at  $t_o$ , or not. If they do have such intentions, and if the intentions generate no conflicts, then their intentions rule:  $o$  is in the extension of  $w$  just in case they intend  $w$  to apply to  $o$ . If they have no such intentions, or if the intentions generate conflict, then (Brown claims) it is most plausible to say that it is indeterminate whether  $o$  is in the extension of  $w$ .

There are variations on the “arthritis” case that remain problematic for such a view. For example, we can imagine a character who shares Arthritis-man’s view about where arthritis can occur, but is himself a medical expert (or self-styled “expert”) with idiosyncratic views. He might be able to offer significant arguments for regarding certain diseases of the thigh as of the same type as the diseases of the joints that are typically called “arthritis”: perhaps they have the same etiology, respond to similar treatments, and so forth. He might have no intention whatever simply to defer to experts. But he is in principle disposed to accept their opinion in certain cases: for example, if they present him with powerful arguments for their view, or with empirical information of which he is unaware. Still, it is by no means a sure thing that he will defer.

In this case, the self-styled expert’s referential dispositions do seem to generate a conflict. On the one hand, he has an intention, grounded in his own (presumed) expertise, to pick out a certain sort of disease in the thigh. On the other hand, he has an independent intention, grounded in his general rationality and respect for the knowledge of his peers, to respond to persuasive arguments. Brown’s view would thus entail that “arthritis” as he uses it is indeterminate in content. But the fact that there could be reasoned disagreement

between Arthritis-man and the experts suggests that they are using the term with the same meaning; and if our character is eventually convinced by some argument, it seems that both of the Burge arguments can easily be applied:

**Response to Correction** He might later go on to think, “I used to think that I had arthritis in my thigh. What a silly mistake!” Such a reflection seems reasonable and correct, and the best explanation of this fact is the hypothesis that Arthritis-man used “arthritis” with the same meaning as the experts all along. In particular, Brown’s view predicts that Arthritis-man’s utterance is untrue: his prior belief was indeterminate in content, and hence not the belief that he had arthritis in his thigh. It is unclear why Arthritis-man would make such an error.

**Disagreement** In light of Arthritis-man’s acceptance of the doctor’s correction, and in light of the fact that he could regard himself as having been mistaken about arthritis all along, it is natural to regard him as agreeing and disagreeing with other speakers even before he is corrected by the doctor. For example, he could agree with a doctor that arthritis is a disease, and in such a case he and the doctor would both believe that arthritis is a disease. This could be true even if the doctor in question is a monolingual speaker of Swahili, with whom Arthritis-man shares no relevant metalinguistic beliefs. Proponents of Brown’s view must deny that Arthritis-man agrees or shares beliefs with the doctor. But this denial is implausible.

Thus Burge’s arguments do not depend in a straightforward way on the existence of an antecedent intention or disposition to defer.

### 2.1.2 Temporal Externalism in Burge’s Cases

Consider again the self-styled expert who is disposed to defer to certain kinds of evidence, and suppose that he also has the following disposition: if after a certain period of debate he is not confronted with evidence he finds convincing, he will conclude that his interlocutors are talking past him, and will be disposed to say things like, “As I’m using the term, ‘arthritis’ can occur in the thigh. You’re just quibbling about the meanings of words; we’ve been talking past each other all along. Sheesh.” This response makes sense given his previous attitudes; it may even be rational, depending on the sorts of arguments to which he is exposed. In this sort of case, the Burge arguments do not apply: it is not plausible to regard him as using “arthritis” with the same meaning as his interlocutors. So in this sort of case, whether the self-styled expert should be regarded as having used “arthritis” with the same meaning as his interlocutors all along depends on how he responds to correction.<sup>16</sup>

<sup>16</sup>What if the correction is never issued, so that the speaker never gets the chance to respond? One natural view would be that in this case, the meaning of the original utterance is indetermi-

There are of course cases in which the subject's earlier attitudes dictate that deferring or refusing to defer (i.e., sticking to one's guns) is correct: perhaps she has a very strong disposition to defer to expertise, or perhaps she has an overriding intention to pick out the disease in her thigh. But the case of the self-styled expert is not like this. The correct response is not fixed at the time of the original utterance. Neither option is dictated by the subjects' dispositions and attitudes at that time; either response is permissible.

Probably few of us have precisely the dispositions of the self-ascribed expert. But broadly similar patterns of dispositions are very common. Though most of us are prepared to defer to experts in some cases, there are many circumstances in which we do not defer blindly: we are prepared to ask for reasons, to argue a point, and (sometimes) to fail to be convinced and to decide that the dispute is merely verbal. Many of these dispositions are rational, and our responses will seem reasonable and correct. But again, our dispositions and other attitudes are complex, and do not determine a unique response at the time of utterance.

Thus temporal externalism is true in familiar cases, and is established by the familiar Burge arguments. (Thus, although temporal externalism is sometimes dismissed as "deeply counterintuitive" (Brown, 2000, p. 178), the patterns of reasoning that can be used in its favour are quite widely accepted.<sup>17</sup> This pattern

---

nate. But the temporal externalist need not endorse this view, at least not in every case. Temporal externalism is a claim about metaphysical dependence. Assume for the sake of argument that it is spelled out in terms of supervenience. Then the temporal externalist claims that any supervenience base for semantic properties must include future-directed properties – as I have developed the view, properties like *accepting later attempts at correction* or *sticking to ones guns in the face of attempted correction*. But this is entirely compatible with the claim that there is determinate content even when these properties are not instantiated. For example, someone might hold that in the open texture cases, that speakers' decisions can make it the case that they used a term with one or another meaning, but that if they make no such decisions (either because the issue never arises, or because the debate is never resolved), some other factor – perhaps Lewisian naturalness – steps in to make content determinate.

<sup>17</sup>Brown presses a further line of argument against the temporal externalist. She claims that TE "fails to accord with our ordinary ways of assessing the truth value of utterances" (p. 186) in at least the following two respects. First, she claims that it does not reflect our epistemic practice with respect to assertion. If TE were true, then speakers would have "an epistemic duty to seek out information about future linguistic practice [...] if it is available" (p. 186). Second, Brown claims that "Speakers would take it that the truth value of [their utterances] depends only on linguistic practice at times up to and including the present [...] not future linguistic practice" (p. 186). On the first point, Brown is correct that we do not typically seek out such information. I take this to be because such information is not typically available. If it is available, it seems very natural to take ourselves to have an epistemic duty to take it in to account. For example, if I use the word "arthritis" to apply to cases of a certain sort, the information that I will soon cease using the word to apply to cases of that sort seems highly relevant. On the second point, Brown is no doubt correct that many speakers would find temporal externalism *prima facie* implausible. But ordinary speakers would might well take the claim that the truth of their utterances depends on

of argument should be of interest even to those who reject my judgments about the cases, as we will apply it again, below.)

I conclude that at least some variations on the Burge scenario shares the structure of the MacFarlane/von Fintel and Gillies cases described above.<sup>18</sup> To put the point explicitly:

**The Data Pattern** There are circumstances under which:

1. A person who does not know that “arthritis” is used in English to refer only to diseases of the joints can reasonably make an assertion using the sentence “I have arthritis in my thigh”.
2. A person who does know how “arthritis” is typically used may respond by rejecting this assertion (e.g., by saying “No, you are wrong”, or “No, you spoke falsely”, or “No, arthritis is a disease of the joints”).
3. The original speaker may either retract her original assertion, or reiterate her original assertion. If she retracts, then this retraction seems correct (and her original assertion seems incorrect). If she reiterates, then her reiteration seems correct (and her original assertion also seems correct).

Moreover, in the variant where Arthritis-man is a (self-styled) expert poised to give reasons for thinking that his disease is arthritis after all, it should be obvious that the Burge scenario shares the structure of the Richard cases.

Given the striking similarity between the Burge cases, the MacFarlane/von Fintel and Gillies cases, and the Richard cases, it is reasonable to look for a common account of all three. I will suggest that temporal externalism puts us in a position to give such an account. In the next two sections, I develop the temporal externalist analysis further. I then turn this analysis to the cases with which we began.

---

the causal history of their words prima facie implausible as well. In my view, the Data Pattern described in the main text shows that TE is a natural explanation of our ordinary ways of assessing at least some utterances, in much the way the cases described by Kripke make it plausible that causal history is in fact relevant to our ordinary ways of assessing utterances involving names.

<sup>18</sup>There is, of course, a difference between the MacFarlane/von Fintel and Gillies cases and the Richard cases on the one hand, and the Burge cases on the other, in that the former but not the latter are concerned with expressions that are relevantly context-sensitive. But this does not seem to bear on the efficacy of the Burge arguments (which I will discuss in more detail with respect to the MacFarlane/von Fintel and Gillies cases in section 3 below). We generally take people who use the same context-sensitive expressions to agree and disagree only when they use those expressions with the same content. For example, you and I do not disagree if I say, “I am hungry” and you say, “I am not hungry”, and I would be irrational if I were to revise my belief that I am hungry on the basis of your utterance. On the temporal externalist view under consideration, the same sorts of metasemantic mechanisms play a role in setting contextual parameters, and in shaping standing meanings.

### 3 Temporal Externalism and the Relativist Data

The externalist account of the “arthritis” scenario, as extended to temporal externalism, gives an account of accommodation and negotiation that is superior to Richard’s. Moreover, it can give an account of the retraction data that motivated MacFarlane, which can also admit von Fintel and Gillies’s “stick to your guns” cases.

First, the MacFarlane retraction case. Recall that the standard semantics for epistemic modals models them as quantifiers over restricted domains of possible worlds. In order to show how the temporal externalist view works, it will be necessary to make some assumptions about how the domain restriction is implemented. On one view (inspired by Kratzer (1977)), modals have syntactically real but typically unpronounced argument places for a proposition called *the modal base*. This proposition is the set of worlds over which the modal quantifies.

The semantic value of the modal base parameter will vary with context. Nonetheless, some metasemantic story must be told about how the modal base for a particular utterance gets fixed. What fixes the modal base? Compare a similar question: what fixes the referent of a particular utterance of “that”? At least three factors might be relevant:<sup>19</sup>

**Lexical meaning** The lexical meaning of “that” does not provide a great deal of information about its referent, but “that” contrasts with “this” in that it tends to indicate objects further from the speaker. (It is a *distal* demonstrative.)

**Environmental cues** A speaker who utters “That’s fuzzy” as she watches a goat walk in is naturally interpreted as referring to the goat. Speakers can generate their own environmental cues, notably by pointing at the object to which they mean to refer.

**Speaker intentions and dispositions** A speaker might intend to use a demonstrative to refer to a particular object, or she might have more complex (e.g. Gricean) reference-related intentions.

A similar set of factors are relevant to determining the modal base:

**Lexical meaning** Certain modal bases are associated with particular modal expressions. For example, “Donna might be delivering Meals on Wheels” cannot normally be used to mean that Donna is permitted to deliver Meals on Wheels because “might” is not normally used to express deontic modality.

---

<sup>19</sup>Some theorists maintain that not all of these factors are actually relevant (e.g. Wettstein (1984); Gauker (2003), see Reimer (1992) for discussion); I am listing them only as candidates.

**Environmental cues** Von Fintel and Gillies (2008, p. 83) have pointed out that in the context of a game of Mastermind (in which one player knows the solution to a puzzle and is meant to give the other player hints that will help her guess), an utterance of “There might be two reds” is naturally interpreted as being true if and only if for all *the hearer* is in a position to know, there are two reds (even if the speaker knows that there are not two reds). In more typical contexts, such an utterance would be false if the speaker knows that there are not two reds.

**Speaker intentions and dispositions** Perhaps most importantly, it is very plausible that speaker intentions and dispositions at the time of utterance play a major role. If a speaker intends her utterance of “The keys might be on the table” to be understood epistemically, and is in the sort of circumstance where she could reasonably expect her audience to pick up this intention, then she has used “might” epistemically. On a simple view of epistemic modals, this will make it the case that the modal base relevant to her utterance consists of the set of worlds compatible with what some agent or agents knows.

In some cases these factors may entirely determine the modal base. For example, a speaker might utter “The keys might be on the table” intending exclusively her own knowledge to be relevant. But in many cases, speakers do not seem to have any such specific intentions. A speaker might utter “The keys might be on the table” without having any explicit or implicit intentions as to whether it is her own knowledge or the knowledge of the entire group that matters. (She might respond to questioning by saying, “I don’t know; the question hadn’t occurred to me”.) She might have some determinate intentions on the matter: perhaps if the question were to occur to her, she would deem irrelevant the knowledge of an omniscient god. But she need not have intentions that distinguish between its being only her own knowledge, or the knowledge of her interlocutors, the knowledge of eavesdroppers, the knowledge of future researchers listening to a recording of her utterance, knowledge contained in nearby information sources (as in Hacking’s (1967) ship’s log case), and so forth. Similarly, her dispositions might be conflicting, confused, or variable, as in the cases discussed above.

What determines the modal base in this sort of case? I will investigate this question in two steps. First, I will look to what the modal base is in some specific cases; then I will ask what might generate a modal base of this sort.

Suppose that an utterance of “The keys might be on the table”, made with indeterminate intentions of the sort described above, is corrected by someone more knowledgeable (“No, it’s not the case that they might be on the table – I already checked there.”) And suppose that the original speaker responds with a MacFarlane-style retraction (“What I said was false. It was not the case that



the keys might be on the table.”)<sup>20</sup> What modal base is relevant to the original utterance? Our evidence seems to be the behaviour of the participants in the conversation. Our account should make sense of this behaviour: in particular, since the contributions of both participants seem perfectly reasonable, our account should make rational sense of their behaviour (to the extent that it is possible to do so).

Now recall Burge’s arguments. Burge suggested that the fact that Arthritisman agrees and disagrees with more expert speakers, and the fact that he accepts the correction of his doctor, are best explained by the hypothesis that Arthritisman uses “arthritis” with the same meaning as the experts, and has the same concept as the experts. In the MacFarlane scenario, the original speaker disagrees with a more knowledgeable speaker, and accepts that speaker’s correction. There is a difference between the “arthritis” case and the case of epistemic modals in that what is at issue in the former is the standing meaning of “arthritis”, while what is at issue in the latter is the semantic value of a parameter whose value is determined by context. But the same styles of argument can be applied in both cases.

On the simplified Kratzerian semantics under consideration, an utterance of “The keys might be on the table” will express a proposition that is true just in case there is some world in the contextually supplied modal base B in which the keys are on the table. Setting aside the context-sensitivity of the definite descriptions “the keys” and “the table”, two utterances of “The keys might be on the table” will agree just in case they are made in contexts that determine the same modal base. Similarly, an utterance of “The keys might be on the table” disagrees with an utterance of “It is not the case that the keys might be on the table” if they are made in contexts that determine the same modal base. (Arguably, they also disagree if the modal base determined in the former context is a subset of the modal base determined in the latter.) Given this model, Burge’s discussion suggests two possible accounts of the MacFarlane scenario:

1. The modal base is the same for the original utterance, the correction, and the original speaker’s retraction.
2. The modal base is not the same for all three utterances. The speakers’ behaviour is explained by false metalinguistic beliefs to the effect that the correction “It is not the case that the keys are on the table” expresses the

---

<sup>20</sup>In order to make the discussion more explicit, I am considering a case in which the original utterance, the attempted correction, and the retraction all involve uses of epistemic modals. But this is inessential: the correction could just be, “You’re wrong”, and the retraction, “Okay, then I take it back”; on the temporal externalist view, this sort of retraction too can play a role in determining the meaning of the original utterance.

negation of the proposition expressed by the original utterance of “The keys might be on the table.”

As in the “arthritis” case, the second hypothesis looks implausible. The hypothesis of false metalinguistic beliefs might explain why the participants in the dialogues take themselves to disagree. But it does not explain why *we* take them to disagree. It seems true that they disagree over whether the keys might be on the table; the one believes that they might, while the other denies this. Moreover, such false beliefs themselves call out for explanation. Why would speakers be so persistently confused in this sort of situation?

Relativists agree with this diagnosis of the case. But what of the von Fintel and Gillies “stick to your guns” cases? Consider the case of a speaker much like Arthritis-man who is unwilling to accept the correction of a doctor and unwilling to negotiate. In this case, Burge’s arguments get no grip. One cannot argue from agreement and disagreement because the Arthritis character who does not accept correction is not naturally seen as agreeing and disagreeing with the doctor (except over the word “arthritis”), and one cannot argue from the way the character responds to correction, because the best explanation of this response seems to be that the character is not using the word “arthritis” to mean *arthritis*. Similarly, in von Fintel and Gillies’s cases, the Burge arguments give us no reason to postulate a modal base common to the original utterance of “It might be...” and the reply of “No, it can’t be...” On the contrary, the apparent rationality and correctness of the “stick to your guns” reply are better explained by the hypothesis that there is no common modal base.

If this is all right, then our theory must account for three facts:

1. In the MacFarlane retraction case, the modal base is the same in the original utterance, the correction, and the retraction.
2. In the von Fintel/Gillies case, the modal base differs between the original utterance and the correction.
3. The speaker need not have any determinate intentions at the time of utterance as to whose knowledge is relevant to determining the modal base, and in particular as to whether (or in response to whom) she will retract or stick to her guns.

If we insist that the modal base must be determined by the intentions of the speaker *at the time of utterance*, these three facts cannot all be accommodated. But if we are temporal externalists, then the three can be accommodated without difficulty, since the speaker’s intentions *at the time of correction* can be relevant to determining the modal base *of the original utterance*. If the speaker accepts the correction, this *makes it the case* that the modal base of her original utterance

includes only worlds compatible with what the corrector knows at the time of correction. If, on the other hand, the speaker sticks to her guns, this makes it the case that the corrector's knowledge is irrelevant to the modal base. In either case, the speaker's response is rational and correct.

I pointed out above that the seeming correctness of objections by eavesdroppers tells against simple contextualist views, since it is implausible that speakers generally intend the knowledge of eavesdroppers to be relevant to the truth of their utterances. Relativist views are designed to handle the cases neatly. The relativist predicts that any eavesdropper can evaluate a given proposition at her own context; if she knows more than the speaker, then the relativist claims that it is perfectly in order for her to deny an epistemic "might" claim. Can the temporal externalist do better? First, note that if an eavesdropper inserts herself into the conversation, other participants are free either to retract or to stick to their guns. Suppose I say in the cafe, "Dilip might be in Vegas," and a stranger at the next table responds, "No, he can't be – I just saw him in the pub." I could then retract or stick to my guns in the normal way. If I stick to my guns, then I have made it the case that the stranger's knowledge is irrelevant to my original utterance, and hence that the stranger is talking past me.

But this fact shows that the relativist's prediction that eavesdroppers can always successfully deny overheard epistemically modalised utterances is incorrect. After all, it could turn out that the content of the utterance is such that only the original utterer's knowledge is relevant. In this case, the eavesdropper would be talking past the original utterance, and failing genuinely to disagree. Until the content of an epistemically modalised utterance is made precise, we should maintain that it is indeterminate whether the eavesdropper's denial is appropriate.

Temporal externalism thus makes it possible to account for both MacFarlane's cases and von Fintel and Gillies's cases, and it does so while maintaining standard views about the semantics of modals. But the case of epistemic modals is an especially simple one. In general, the modal base will steadily become more and more precise over the course of a conversation; it is hard to imagine much substantial negotiation (giving arguments, offering reasons for one view or another) of the sort that can occur in other sorts of case. (This explains the fact that after the utterer of an epistemically modalised sentence has responded to attempted correction by sticking to her guns, further debate is usually impossible.) But the same metasemantic theoretical machinery can explain other sorts of cases where it is possible to give further arguments and reasons. Richard described two people arguing about whether Didi is rich, despite agreeing on the relevant comparison class and the facts Didi's level of income. Roughly speaking, the parties to the dispute are applying different criteria: one thinks that to be rich for a New Yorker is to make more than \$300,000 per year, while the

other maintains that to be rich for a New Yorker is to make more than \$600,000. Richard argues that such debates play a role in shaping our concepts.

What is needed is a view that (like relativism) gives an object of disagreement (where one is needed), and (unlike relativism) makes sense the idea that the result of such debates could be conceptual change. And there is a further datum to be considered. Suppose that one of Richard's debaters manages to convince the other that Didi is indeed rich. Then it will seem to the debaters that their debate was genuine: they were not simply talking past each other. They would be inclined to report, "We used to disagree about whether Didi is rich, but now we agree", and it is very natural to regard this report as true. But consider, on the other hand, the case where neither party convinces the other, and the debate simply ends. This could happen in two ways. First, the parties could end the debate because (for example) they are out of time, though they acknowledge that there are more reasons to be given on either side and the argument has not reached a natural stopping place. (Typically, in this sort of case, they might be willing to continue the debate at a later time.) But second, the parties could regard the reasons on both sides as being exhausted, and could regard the result as an irresolvable stalemate, where there is no more to be said. In this second sort of case, we are much more likely to regard the parties as talking past each other. It is much more plausible in this case to regard the debate as merely verbal – not in any robust sense a debate about whether Didi is rich at all, but a debate about whether the sentence "Didi is rich" is true.

In short, the situation in Richard's cases is as follows. If you say "Didi is rich", I am faced with a choice: I can accommodate the utterance, go along with this way of speaking, and use "rich" (for the purposes of the conversation) to mean something like (say) "rich enough to eat in fancy restaurants in New York"; or I can refuse to accommodate and offer a correction (thus entering into negotiation). Similarly, when faced with my "correction", you face a choice: you can accept my usage, or you can refuse. If you refuse, you can continue to negotiate by offering reasons to think that "rich" as you use it is a useful way to categorise.<sup>21</sup> If at any point one party refuses to accommodate and refuses to negotiate then (*ceteris paribus*) we should not treat the two as using "rich" with the same meaning.<sup>22</sup>

---

<sup>21</sup>Of course, we should not over-intellectualise the process of negotiation. In real conversations, we are not always able to respond to every consideration against a view. I may not find your argument compelling even if I cannot say exactly what is wrong with it, and my capacity to defend my own view may be limited. Such a failure to engage with the considerations advanced by an opponent does not in itself constitute a refusal to negotiate. We can regard the conversation as ongoing, the issue as unresolved, and ourselves as continuing to disagree, even if at some point we find ourselves with nothing more to say. Refusal to negotiate consists at least in part in a different set of attitudes: regarding the argument as over, and ourselves as failing to disagree. I return to this point in section 3 below.

<sup>22</sup>"*Ceteris paribus*" because there are cases in which we might refuse to accommodate and

In that case, there is no genuine disagreement; the parties are really talking past each other.

If this is right, then one major motivation for Richard's relativism is undermined. Richard wants to reconcile the claim that the arguers genuinely disagree with the claim that each speaker can correctly regard his pre-argument utterances of "Didi is rich" and "Didi is not rich" (respectively) as true. But this motivation is confused. It is true that it might turn out that the arguers genuinely disagree (if one convinces the other), and it is true that both pre-argument utterances might be true (if the argument ends with no accommodation and a refusal to negotiate further). But in the latter case, there will have been no genuine disagreement, even by the arguers' lights; and in the former case, one or the other arguer will admit that his original utterance was untrue. So plausibly there is no case of the sort Richard wants to account for.

All of these data are exactly what the temporal externalist predicts. If one party ends up "deferring" to the other's "expertise", then the case is like Burge's: this deference plays a role in making it the case that both parties to the debate are meaning the same things, are genuinely disagreeing, and so forth. So in this case, there is an object of disagreement: the (possible worlds, non-relativised) proposition (e.g.) that Didi is rich – though one party begins the debate with a false belief about this proposition. The parties' dispositions to apply the term at issue at the end of the debate help determine the content of the term as used at the beginning of the debate. So the idea that this sort of case involves change in meaning or conceptual change is also accommodated. If, on the other hand, a stalemate is reached with neither party conceding, then the case is much more like the case in which Arthritis-man fails to defer to the doctor: in both, the best explanation of the data seems to be that the two use the word with distinct meanings and have distinct concepts. So the temporal externalist also predicts that the stalemate case is one of merely verbal disagreement.<sup>23</sup>

---

refuse to negotiate for practical reasons (e.g., our meeting is about to end and I have to catch a bus), or because we fail to see how make further progress and regard pursuing the matter as futile (and so "agree to disagree"). In at least some such cases, it is very plausible that we could genuinely disagree. But these cases are not the sort of final refusal to continue that I have in mind; here the refusal amounts to an (at least in principle) temporary suspension of negotiation, rather than an ultimate termination.

<sup>23</sup>Some might see another problem for the temporal externalist who also wants to oppose relativism. Relativism has been suggested as a response to the open future (MacFarlane, 2003, 2008): if the future is open, then it seems that an occurrence of "There will be a sea battle at  $t_0$ " might be not determinately true as evaluated before  $t_0$ , but determinately true as evaluated at a later time  $t_1$ . In particular, supposing that there turned out to be a sea battle at  $t_0$ , I could say at  $t_1$ , "What I said before was true, and I stand by it"; while if there turned out to be no sea battle at  $t_0$ , it would be appropriate for me to say at  $t_1$ , "What I said before was false, and I retract it". MacFarlane argues that this can be explained on the relativist hypothesis that future contingents

## 4 Temporal Externalism and Assertion

Temporal externalism explains a great deal of otherwise troublesome data. But a puzzle remains. When someone utters “Donna might be delivering Meals on Wheels” or “I have arthritis in my thigh” with suitably indeterminate intentions and dispositions, then what proposition they express depends on the future. Now if the future is open, these utterances may not have a determinate semantic value; and even if the facts about the future are fixed, we may not be in a position to know the facts that determine what the semantic value is. Plausibly, there is even a sense in which we may not know what proposition is asserted. So what are we doing when we make an assertion, if temporal externalism is true?

I want to argue that this problem can be resolved with tools that are needed by every account of assertion to deal with cases in which an audience is in some sense not in a position to know what proposition is expressed by an utterance. Temporal externalism fits neatly with two different approaches to assertion: Stalnaker’s model, which treats assertion in terms of its effects on the context, and Williamson’s account, which treats assertion in terms of norms. My strategy will be as follows. First, I will present a case that Stalnaker (1978) uses to motivate *diagonalisation*, and will discuss Stalnaker’s treatment of the case. I will then show that the machinery developed to handle Stalnaker’s case – beginning with Stalnakerian diagonalisation and returning later to an alternative framework – can be adapted to treat assertion in a temporal externalist framework. This discussion will make it clear how the temporal externalist can respect Williamsonian norms of assertion. So the temporal externalist can account for assertion using tools needed to handle familiar cases, cases that any theory must explain.

### 4.0.3 The Stalnakerian Account

On Stalnaker’s view, the crucial fact about an assertion is that it has certain effects on the context in which it is made. Stalnaker models propositions as sets

---

are assessment sensitive. Similarly, if the future is open, the temporal externalist seems to predict that an utterance of “What you said means that for all you know, Donna is delivering Meals on Wheels” might be not determinately true as evaluated at one time, but determinately true as evaluated from a later time (after the facts that determine the meaning are settled). A similar line of argument would suggest that the truth of claims about meaning is assessment sensitive.

As I pointed out in footnote 3, this motivation for relativism is different than the sorts of motivations that are the main focus of this paper, in that it is based on a controversial claim about the metaphysics of time. Temporal externalism does not undermine this metaphysical motivation for relativism; if the future is open, and if this fact can be used to defend relativism, temporal externalism would make claims about meaning relatively true in just the way claims about sea battles are. In my view, this kind of relativist argument can be answered by careful consideration of the metaphysics of time, but this raises difficult issues that I cannot consider here. Even if relativism could be defended in this way, the temporal externalist account of the sort of data that is alleged to motivate relativism in the other cases would remain significant (especially if the arguments against relativism in sections 1.2 and 1.3 above succeed).

of possible worlds, and a context as a set of propositions – the propositions that the participants in the conversation presuppose for the purposes of the conversation. Making an assertion will typically have many effects on the context; for example, it will typically be the case that when I make an assertion, all of the participants in the context will come to presuppose that I have spoken. But one of these effects is particularly important on Stalnaker’s view: an assertion of a proposition aims to modify the context by removing those worlds from the context that are incompatible with that proposition. This is the *essential effect of an assertion* (1978, p. 86).

The following situation presents two problems for this picture of assertion: “hearing a woman talking in the next room, I tell you, *That is either Zsa Zsa Gabor or Elizabeth Anscombe*” (1978, p. 91). The first problem arises specifically within Stalnaker’s framework. Given that demonstratives and names are rigid designators, the sentence “That is either Zsa Zsa Gabor or Elizabeth Anscombe” expresses a proposition that is necessarily true if true at all. This is problematic given Stalnaker’s view that propositions are sets of possible worlds, since on this view, a necessarily true proposition conveys no information. Since the utterance described could be informative, Stalnaker needs to tell a story about how some contingent proposition comes to be expressed by the utterance. The second, more general, problem is that there is a sense in which such an assertion makes sense only if the speaker and the audience are not in a position to know what proposition is asserted. (Even though plausibly there is a sense in which they are in a position to know what proposition is asserted – it is the proposition that that is either Zsa Zsa Gabor or Elizabeth Anscombe – Stalnaker insists that there is another sense in which they do not know – for example, they do not know whether it is the proposition that Zsa Zsa Gabor is either Zsa Zsa Gabor or Elizabeth Anscombe. I return to the denial of this Stalnakerian claim below.)

Stalnaker begins by considering the fact that this utterance semantically expresses different propositions in different worlds: if Gabor is the woman talking in the next room (and hence the referent of “That”), then the utterance expresses a proposition that is true just in case Gabor is either Gabor or Anscombe; if Saul Kripke is in the next room, then the utterance expresses a proposition that is true just in case Kripke is either Gabor or Anscombe; and so forth. If we choose to accept the assertion, how should we update our beliefs and presuppositions? If the actual world is a world in which Gabor is in the next room, then the utterance expressed the proposition that Gabor is either Gabor or Anscombe, and that proposition is true. If, on the other hand, the actual world is one in which Kripke is in the next room, then the utterance expresses a false proposition. So if the utterance is true, we know that the actual world is not one in which Kripke is in the next room. And more generally, we know that it must be either Gabor or Anscombe in the next room, since these are the only possibilities (compatible

with what we are taking for granted about the semantics of English) on which the utterance expresses a truth. Stalnaker concludes that this is what is asserted by his utterance; roughly, it is the proposition that the proposition semantically expressed by his utterance is true (1978, p. 82).

This is Stalnaker's *diagonalisation* technique. According to Stalnaker, the diagonal is *what is asserted* by an utterance like the one described; the essential effect of the utterance is to add the *diagonal* to the context (rather than the proposition semantically expressed by the utterance). But it is also worth noting that in normal cases, the diagonal will be added to the context anyway – not in virtue of the essential function of assertion, but simply as something that participants in the conversation will come to accept in virtue of their linguistic knowledge and the assertion. (After all, it is more or less trivial that most cases in which we accept an utterance are cases in which we accept that proposition semantically expressed by the utterance is true.)

Recall now the second intuitive issue raised by Stalnaker's example: that the most natural context in which such an utterance would take place is one in which the speaker and the audience are in some sense not in a position to know what proposition is semantically expressed. Suppose, for example, that Stalnaker continues his utterance by saying, "She sure knows a lot about modal logic!" This expresses either a proposition that is true just in case Gabor knows a lot about modal logic, or a proposition that is true just in case Anscombe knows a lot about modal logic. But if speakers do not know which, they do not know which possibilities are compatible with what is said. And given Stalnaker's possible-worlds model of belief and presupposition, this means that they will not be able to update their attitudes appropriately.

On Stalnaker's view, diagonalisation helps here as well. How should we respond to the assertion? Should we update our beliefs? If so, how? Well, one thing we know is that if the assertion is correct, the diagonal is true: i.e., the actual world is either one in which the person in the next room is Gabor and Gabor knows modal logic, or one in which the person in the next room is Anscombe and Anscombe knows modal logic. This is a proposition that we understand. If we take the speaker to be authoritative, it would make sense to update our beliefs with this proposition. And even if we doubt that the speaker is right, we can use the diagonal to proceed with the conversation: it something we can accept for the purposes of the conversation, or choose to dispute.

Since the problem raised by the temporal externalist cases is so similar to the problem raised by the Anscombe/Gabor case, it is natural for the temporal externalist to appeal to diagonalisation. It is perhaps unlikely in this case that we would take the speaker to be authoritative, but if we do, we will rule out worlds in which she falsely says that arthritis can occur in the thigh; for example, we can rule out worlds in which she will defer and experts use "arthritis" to mean



*arthritis*. So if we accept the assertion, we will come to believe that either she will defer and experts use “arthritis” to apply to (inter alia) diseases of the thigh, or that she will not defer and is using “arthritis” (possibly idiosyncratically) to apply to diseases of the thigh.

If we choose not to accept an assertion, we have several options. Consider a simple case: you say “Zsa Zsa knows about modal logic.” I think that I understand what you said. But I also know that although Zsa Zsa is an expert in category theory, she does not know about modal logic. Various hypotheses suggest themselves:

**Possibility A** You might be using words differently, so that I have not really understood what you said. For example, you might be speaking of a different Zsa Zsa, or you might be using the term “modal logic” to refer to category theory.

**Possibility B** You might be misinformed about Zsa Zsa’s competencies.

Now if I know that Possibility A obtains, I can diagonalise. But if Possibility B obtains, I don’t want to believe what you said, or even the diagonal of what you said (since these will both be false). If I know that the latter case obtains, my most likely response will be to deny what you say, for example by replying, “No, it is category theory that she knows.” (Of course, there are cases in which it would be rude or imprudent to confront you in this way; if I am suitably motivated to avoid conflict, I might go along with you for the purposes of the conversation. In that case, I would not come to believe what you assert, but I would come to presuppose it, if only temporarily.)

Now suppose that I do not know whether Possibility A or Possibility B obtains. Then I will not know how I should respond to your utterance. If Possibility A obtains, I should believe the diagonal; if Possibility B is true, then I should deny what you said. There are various strategies I might adopt, depending on the situation. If I am feeling confrontational, I might deny what you said: the resulting conversation is likely to clarify the situation. If I am intimidated by you, I might simply make as if I have accepted what you said, and let the unwinding conversation reveal wherein your mistake or our misunderstanding lies. Or I might ask for clarification as to what you meant, or why you think what you do.

Exactly the same possibilities arise in the temporal externalist cases we have been focusing on. As Burge pointed out, when confronted with an utterance of “Arthritis occurs in the thigh”, we can deny what was said; the resulting conversation is likely to make it clear whether the speaker is using “arthritis” as the experts do. But we could equally play along, or, perhaps more likely, ask what the speaker meant. And as MacFarlane pointed out, we can respond to utterances of “The keys might be on the table” by denying what was said. But we can

also (if a bit fussily) ask whether the speaker meant that the keys might be on the table *for all she knows*.

#### 4.1 Alternatives to Diagonalisation

I have so far developed the view of assertion by appeal to diagonalisation. But the claim that diagonalisation has an important role to play in the theory of communication is highly controversial. Still, everyone needs to give some account of utterances like Stalnaker's "She knows a lot about modal logic," and there is no reason to think that such accounts are likely to be incompatible with temporal externalism. For example, Hawthorne and Magidor suggest that one can appropriately update one's beliefs in response to an utterance by simply accepting the uttered sentence (2009, p. 394). One might not be in a position to recognise that proposition under a different "guise": for example, as the proposition that Zsa Zsa knows about modal logic. One might therefore be unable to recognise certain inferences as valid, or to recognise the conflict between this proposition and one's antecedent beliefs. But on this view, this does not prevent one from accepting the proposition.

Once we have taken on board the idea that we can hold beliefs by accepting sentences even in cases where we cannot recognise the proposition expressed by the sentence under other guises, then there seems to be no bar to the idea that we can hold beliefs by accepting sentences like "Arthritis sometimes occurs in the thigh", when we do not know how the utterer will react to correction. It is of course true that this will limit what one could do with such a belief; for example, one might not be able to recognise certain valid inferences that the belief would make possible, or to recognise conflicts with one's other beliefs. But these are consequences that a proponent of this style of view should already accept.

#### 4.2 Norms of Assertion

All this is suggestive. But there is a problem which prevents the temporal externalist from maintaining that the diagonal is what is asserted. It is plausible that assertion is governed by norms; for example, perhaps one must assert only that which one knows, or only that which is true.<sup>24</sup> But if we assert the diagonal in the temporal externalist cases, we are often asserting propositions that turn out to be false, and hence propositions that we do not know. But in the temporal externalist cases, speakers do not seem to be violating norms of assertion in this way. And this style of objection has some force even if we deny the Williamsonian view that the knowledge norm is constitutive of assertion; on any view, it

---

<sup>24</sup>The knowledge norm is championed in chapter 11 of Williamson (2000). For a survey of the post-Williamson literature on this topic, see the introduction to Brown and Cappelen (2011) and the works cited therein.

is plausible that assertions are made irresponsibly if the asserter is in a position to know that they may well be false. But it may seem that we are often in such a position, according to the temporal externalist. Yet most of our assertions do not seem irresponsible in this way.

Despite this, temporal externalism is compatible with the knowledge norm of assertion. To see this, first note that we do seem responsive to something like these norms even in the temporal externalist cases. One of the key motivations for relativists like MacFarlane and Richard is that we *retract* (e.g.) epistemically modalised claims or claims involving gradable adjectives when faced with (seemingly) counterevidence. On my view, when we retract (other things being equal), we make it the case that the proposition we asserted was one that we did not know (e.g., that for all we or any eavesdroppers knew, the keys were on the table). But we also retract the assertion: an appropriate response, given that the norm was violated. And similarly, when we stick to our guns we make the case the proposition we asserted is one that we did know (e.g., that for all *I* knew, the keys were on the table). Here there is no violation of the norm, so the reassertion is appropriate. In either case, the course of the conversation will be something quite consistent with a truth or knowledge norm.

How can we reconcile these facts with the use of diagonalisation? Here it is important to distinguish the two ways that the diagonal might get added to the context: either as the essential effect of the assertion (in which case the diagonal is what is asserted), or as a secondary effect of the assertion (in which case the diagonal is not what is asserted). Stalnaker claimed that in the Gabor/Anscombe case, the diagonal is what is asserted. Whether or not this is correct in the Gabor/Anscombe case, it is something that the temporal externalist should deny in the cases that motivate her view; if the temporal externalist claims that diagonal is what is asserted in these cases, then she will predict widespread violations of the norms of assertion where no such violations are apparent. Instead, the temporal externalist should maintain that what is asserted in these cases is the *semantic* content, and this depends upon the future.<sup>25</sup> Participants in the conversation may not know what this content is; indeed, there may be no fact of the matter. So the assertion cannot fulfil its essential function at the time it is made. (Or at

---

<sup>25</sup>The situation will be complicated in a Stalnakerian framework in cases where the semantic content is necessary. For example, suppose I say, "Arthritis just is swelling of the joints or the thigh," intending this to be an analytic or definitional truth. And suppose that if I am corrected by a doctor, I may retract the utterance. The Stalnakerian predicts that the diagonal will be what is asserted; but I do not know the diagonal to be true, because I cannot rule out scenarios in which I am corrected by a doctor and defer. If I assert the diagonal without knowing it, then I have violated the knowledge norm of assertion. My own view is that even in this sort of case, the diagonal is not what is asserted; there is no obstacle (at least in this sort of case) to the idea that a necessary truth is what is asserted once it is realised that the diagonal is added to the context as a secondary effect of the assertion.

least not completely; some worlds might be determinately ruled out.) This is a departure from Stalnaker. But there is no problem about understanding conversations involving such assertions: the assertions have clear effects on the context, since they add the diagonal to it. It is therefore clear how the conversations can proceed.

Now this does yet not go all the way toward answering the original objection. After all, if the semantic content is what is asserted, it is still the case that what is asserted might turn out to be false. Is this not still a violation of the knowledge norm of assertion? Here the crucial fact is that whether or not the assertion violates the norm is within the speaker's control; it is up to her whether or not to accept the arguments of her conversational partners and retract, and hence up to her whether or not to make it the case that her assertion expresses something she knows to be false. The situation is thus rather like making a promise that it is in one's power to keep. It might turn out that one does not keep one's promise, but it is only when one actually fails that one is blameworthy; one cannot be blamed at the time of making the promise merely because one might fail to keep it. Similarly, when one makes an assertion, even if at the time of utterance it is not determinate whether one has asserted something known, one is not blameworthy for having made the assertion until it is determinate that one has asserted something unknown, since until that point it is in one's power to make it the case that one has asserted something known.

But the fact remains that one might make it the case that one asserts something false, and of course in this sort of case the knowledge norm of assertion would be violated. But this sort of violation is not problematic. Contrast two ways in which a rule or norm can be violated. Say that a rule is *disobeyed* if the violator continues to act as if they were bound by the rule – for example, by making excuses, issuing apologies, or trying to make amends for the violation. A rule is *disregarded* if the violator shows no awareness of ever having been bound by the rule in the first place. No view should predict that the norms of assertion are never disobeyed; in some circumstances there are very good reasons to disobey the norms. What a view should not predict is that the norms of assertion are routinely disregarded. If knowledge is the norm of assertion, it is hard to make sense of the view that all of us routinely assert things that we do not know without any sense of transgression. For example, (as I pointed out in section 1.1 above), the form of contextualism that makes  $\ulcorner$ It might be the case that  $\phi$  $\urcorner$  roughly equivalent to  $\ulcorner$ for all we jointly know,  $\phi$  $\urcorner$  predicts that speakers often assert propositions that they do not know. This fact would not be very damaging to the view if the speakers recognised their mistakes as such and tried to make amends. The really damaging fact for this sort of view is that speakers show no awareness of having done anything wrong. Moreover, we theorists considering the case also show no awareness of the rampant disregarding of norms predicted

by such a view. Temporal externalism, on the other hand, predicts that the norm of assertion is disobeyed, but it never predicts that it is disregarded. The very act that makes it the case that I violate the norm is simultaneously a retraction; every time a speaker violates the norm, she also seeks to make amends.

Let me sum up the discussion of this section. The account of assertion for the temporal externalist has three moving parts:

1. Assertions of sentences with a temporal externalist metasemantics typically contribute the diagonal proposition to the context, but this is a secondary effect (rather than the essential effect) of the assertion. Nonetheless, this gives participants something to accept or to object to, so the conversation can proceed.
2. The temporal externalist predicts that sometimes we make assertions that turn out to be false, and hence that violate plausible norms of assertion. But this is acceptable for two reasons:
  - (a) One is not at fault merely for making assertions using sentences with a temporal externalist metasemantics, since it is in one's power to make it the case that these assertions turn out to be known.
  - (b) Should one choose to make it the case that one's assertion is false, then one has violated a norm of assertion; but since the action by which one makes it the case that the assertion is false is also a retraction of the assertion, this violation is relatively unproblematic. (Temporal externalism does not predict that the norms of assertion are disregarded.)

## 5 Conclusion

I have argued that there is an independently motivated metasemantic account of the putatively relativist friendly data, and that this account also explains the "stick to your guns" data that is problematic for the relativist. The phenomenon exhibited by the relativist's cases is structurally similar to Burge's famous examples, and the two should be treated in the same way.

This style of account has potential for application in a wide range of cases: I suspect that temporal externalist mechanisms can help explain the use of conditionals in Allan Gibbard's (1981) Sly Pete case, the sorts of debate between sceptics and their opponents that Keith DeRose (1995) and Stewart Cohen (1986; 1988) used to motivate contextualism about "knows", disagreements about logic, vagueness, and many others. I leave these cases for future work.

## References

- Brogaard, B. (2008). In defence of a perspectival semantics for 'know'. *Australasian Journal of Philosophy*, 86(3).
- Brogaard, B. (2009). Introduction to relative truth. *Synthese*, 166(2):215–229.
- Brown, J. (2000). Against temporal externalism. *Analysis*, 60:178–188.
- Brown, J. and Cappelen, H., editors (2011). *Assertion: New Philosophical Essays*. Oxford University Press, New York.
- Burge, T. (1979). Individualism and the mental. In Ludlow, P. and Martin, N., editors, *Individualism and Self-Knowledge*, pages 21–84. CSLI, Stanford. Originally published in *Midwest Studies in Philosophy*, 4, 1979.
- Burge, T. (1988). Individualism and self-knowledge. *Journal of Philosophy*, 85:649–663.
- Cappelen, H. and Hawthorne, J. (2009). *Relativism and Mondaic Truth*. Oxford University Press, Oxford.
- Cohen, S. (1986). Knowledge and context. *Journal of Philosophy*, 83:574–583.
- Cohen, S. (1988). How to be a fallibilist. *Philosophical Perspectives*, 2:91–123.
- DeRose, K. (1995). Solving the skeptical problem. *Philosophical Review*, 104:1–52.
- Dowell, J. (2011). A flexible contextualist account of epistemic modals. *Philosophers' Imprint*.
- Egan, A. (2007). Epistemic modals, relativism and assertion. *Philosophical Studies*, 133(1):1–22.
- Egan, A. (2010). Disputing about taste. In Feldman, R. and Warfield, T., editors, *Disagreement*. Oxford University Press, Oxford.
- Egan, A. and Weatherson, B., editors (2011). *Epistemic Modality*. Oxford University Press, Oxford.
- Egan, A., Weatherson, B., and Hawthorne, J. (2005). Epistemic modals in context. In Preyer, G. and Peter, G., editors, *Contextualism in Philosophy*. Oxford University Press, New York.
- Field, H. (2001a). Indeterminacy, degree of belief, and excluded middle. In Field (2001c), pages 278–306.

- Field, H. (2001b). Theory change and the indeterminacy of reference. In Field (2001c), pages 177–193.
- Field, H. (2001c). *Truth and the Absence of Fact*. Oxford University Press, New York.
- García-Carpintero, M. and Kölbel, M., editors (2008). *Relative Truth*. Oxford University Press, New York.
- Gauker, C. (2003). *Words without Meaning*. MIT Press, Cambridge, MA.
- Gibbard, A. (1981). Two recent theories of conditionals. In Harper, W., Stalnaker, R., and Pearce, G., editors, *Ifs*, pages 211–247. Reidel, Dordrecht.
- Glanzberg, M. (2007). Context, content, and relativism. *Philosophical Studies*, 136:1–29.
- Hacking, I. (1967). Possibility. *Philosophical Review*, 76:143–168.
- Hawthorne, J. and Magidor, O. (2009). Assertion, context, and epistemic accessibility. *Mind*, 119:1067–1086.
- Jackman, H. (1999). We live forwards but understand backwards: Linguistic practices and future behavior. *Pacific Philosophical Quarterly*, 80:157–177.
- Jackman, H. (2005). Temporal externalism, deference, and our ordinary linguistic practice. *Pacific Philosophical Quarterly*, 86:365–380.
- Kaplan, D. (1989). Afterthoughts. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes From Kaplan*, pages 565–614. Oxford University Press, New York.
- Kaplan, D. (1996). Dthat. In Martinich, A., editor, *The Philosophy of Language: Third Edition*, pages 292–305. Oxford, New York.
- Kölbel, M. (2002). *Truth Without Objectivity*. Routledge, New York.
- Kölbel, M. (2004). Faultless disagreement. *Proceedings of the Aristotelian Society*, 104(1):53–73.
- Kratzer, A. (1977). What ‘must’ and ‘can’ must and can mean. *Linguistics and Philosophy*, 1(3):337–355.
- Laserson, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, 28(6):643–686.

- Lasersohn, P. (2008). Quantification and perspective in relativist semantics. *Philosophical Perspectives*, 22(1):305–337.
- Lasersohn, P. (2009). Relative truth, speaker commitment, and control of implicit arguments. *Synthese*, 166(2):359–374.
- Lewis, D. (1979). Scorekeeping in a language game. In *Philosophical Papers Volume I*, pages 233–249. Oxford University Press, New York.
- MacFarlane, J. (2003). Future contingents and relative truth. *Philosophical Quarterly*, 53:321–336.
- MacFarlane, J. (2005). Making sense of relative truth. *Proceedings of the Aristotelian Society*, 105(3):321–339.
- MacFarlane, J. (2008). Truth in the garden of forking paths. In García-Carpintero and Kölbel (2008), pages 81–102.
- MacFarlane, J. (2009). Nonindexical contextualism. *Synthese*, 166:231–250.
- MacFarlane, J. (2011a). Epistemic modals are assessment-sensitive. In Egan and Weatherson (2011).
- MacFarlane, J. (2011b). Simplicity made difficult. *Philosophical Studies*, 156:441–448.
- MacFarlane, J. (2014). *Assessment Sensitivity: Relative Truth and its Applications*. Oxford University Press, Oxford.
- Reimer, M. (1992). Three views of demonstrative reference. *Synthese*, 93:373–402.
- Richard, M. (2004). Contextualism and relativism. *Philosophical Studies*, 119(1–2):215–242.
- Richard, M. (2008). *When Truth Gives Out*. Oxford University Press, Oxford.
- Rosen, G. (2010). Metaphysical dependence: Grounding and reduction. In Hale, B. and Hoffman, A., editors, *Modality*, pages 109–136. Oxford University Press, Oxford.
- Schaffer, J. (2009). On what grounds what. In Chalmers, D., Manley, D., and Wassermann, R., editors, *Metametaphysics*, pages 347–383. Oxford University Press, New York.



- Stalnaker, R. C. (1978). Assertion. In *Context and Content: Essays on Intentionality in Speech and Thought*, pages 78–95. Oxford University Press, New York.
- Stalnaker, R. C. (1984). *Inquiry*. The MIT Press, Cambridge, MA.
- Stephenson, T. (2007). Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy*, 30(4):487–525.
- von Fintel, K. and Gillies, A. S. (2008). CIA leaks. *Philosophical Review*, 117:77–98.
- von Fintel, K. and Gillies, A. S. (2011). ‘Might’ made right. In Egan and Weatherson (2011), pages 108–130.
- Waismann, F. (1945). Verifiability. *Proceedings of the Aristotelian Society, Supplementary Volume*, 19.
- Weatherson, B. (2009). Conditionals and indexical relativism. *Synthese*, 166(2):333–357.
- Wettstein, H. K. (1984). How to bridge the gap between meaning and reference. *Synthese*, 58:63–84.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press, Oxford.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwell, Oxford.
- Wittgenstein, L. (1967). *Remarks on the Foundations of Mathematics*. Basil Blackwell, Oxford.
- Wittgenstein, L. (1976). *Wittgenstein’s Lectures on the Foundations of Mathematics: Cambridge, 1939*. University of Chicago Press, Chicago.
- Wright, C. (2008). Relativism about truth itself: Haphazard thoughts about the very idea. In García-Carpintero and Kölbel (2008), pages 157–186.
- Zimmerman, A. Z. (2007). Against relativism. *Philosophical Studies*, 133:313–348.