CONFLICT AND CONTENT

by

Timothy R. Sundell

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Philosophy)
in The University of Michigan
2009

Doctoral Committee:

      Assistant Professor Anthony S. Gillies, Co-Chair
      Professor Peter Ludlow, Northwestern University, Co-Chair
      Professor Samuel D. Epstein
      Professor Peter A. Railton
      Assistant Professor Andrew M. Egan

Contents

Abstract

Speakers differ from one another in philosophically problematic ways. Two speakers can vary not simply with respect to what they believe, but also in the ways they speak, the concepts they employ, and the standards they bring to bear. The fact of imperfect convergence gives rise to a wide range of philosophical puzzles, largely via a single generalization: If two speakers disagree with each other, then at least one of them says something false. The generalization is plausible, but mistaken. Counterexamples are common, diverse, and thoroughly entrenched in ordinary talk, scientific discourse, and philosophical inquiry. I focus on a particular family of counterexamples, disagreements about the proper application of linguistic items, or what I call *metalinguistic disagreements*.

Coming to grips with the widespread existence of metalinguistic disagreement requires a nuanced account of the ways in which information can be transmitted via an utterance, and it suggests a substantial rethinking of conventional philosophical wisdom about the nature of meaning. I begin with a philosophical case study in metalinguistic disagreement—disagreements about aesthetics. I go on to explore the more general linguistic properties of metalinguistic disagreements, focusing on their striking failure to license so-called *metalinguistic negation*. I conclude with a consideration of semantic methodology, arguing that, despite widely held assumptions to the contrary, metalinguistic disagreements can be adjudicated with reference to objective, non-linguistic features of the world. It is possible therefore to roundly reject a pernicious relativism about scientific discourse while simultaneously allowing for widespread, interpersonal variation in meaning.

Chapter 1

Disagreements About Taste

My topic will be a new version of an old puzzle. In its most basic and general form, the puzzle is this: On the one hand, there is no disputing taste. On the other hand, there *is* disputing taste. The version of the puzzle I'll consider goes like this: Consider two speakers, Alphie and Betty. Alphie utters sentence (1a). Betty utters sentence (1b).[1]

    (1)   (a)   Eggo Waffle Cereal is delicious.

              (b)   Nuh uh, Eggo Waffle Cereal is not delicious.

Two very intuitive ideas about dialogue (1) are in conflict with each other. On the one hand, it seems possible that neither Alphie nor Betty is mistaken. On the other hand, it seems that Alphie and Betty disagree.

This version of the puzzle has recently been at the center of debates about the semantics of aesthetic terminology, or as that terminology is sometimes known in that literature, predicates of personal taste. But the tension between no-fault aesthetic variation on the one hand, and the possibility of intersubjective evaluation on the other, has hardly gone unnoticed by philosophers. As will become clear, the linguistic version of the puzzle I'll consider here is importantly different from its purely philosophical cousins, not least because the domain has shifted from aesthetic judgements to the language used to express those judgements. Nevertheless, the broader tension between the subjective nature of aesthetic taste and the potential for substantive and even

---

[1] Following Lasersohn (2005), Stevenson (2007), Egan (2007), MacFarlane (2007), and others, I will focus on words like *delicious* rather than *beautiful* in order to set aside for the moment as much philosophical baggage as possible. I leave open the possibility that there is a great deal of heterogeneity within the class of aesthetic predicates, and that *beautiful* may be quite different from *delicious* in other important respects.

resolvable aesthetic dispute is worth bearing in mind, and I will return to the broader philosophical issues at the end of the paper.

The paper will be in six sections. First, I will lay out in detail the recent version of the puzzle and its relevance to theories about predicates of personal taste. Second, I will distinguish two sets of disagreement phenomena and offer a stipulative definition of disagreement that will allow for a more precise discussion of both. Third, I will describe very briefly the range of situations that can be classed as disagreements. Fourth, I will consider a simple but widespread disagreement-based argument against contextualism about predicates of personal taste. Fifth, I will consider two less common but more sophisticated disagreement-based arguments against contextualism. Finally, in section six I will return to the more general aesthetic puzzle and consider which questions have and which have not been advanced by progress made on the linguistic front.

1. The Puzzle

Recall that the puzzle about dialogue (1) had to do with two conflicting intuitions. Again those intuitions are that, on the one hand, it seems possible that neither Alphie nor Betty is mistaken, and on the other, it seems that they disagree. What do those intuitions have to do with a theory of meaning? Consider the first intuition: that it is possible that neither Alphie nor Betty is mistaken. If we think that someone is not mistaken in uttering a sentence, then the most natural conclusion is that the person has spoken truly. Of course, that inference must be qualified. After all, there are a number of reasons why a speaker might blamelessly utter something false. She could engage in a blameless form of insincerity, by playing a game of pretend, for example. She could attempt to communicate some proposition that contradicts the literal content of her expression, by being, for example, sarcastic. Or she might simply utter something false that she had very good reason to believe was true. Nevertheless, it seems possible that none of those circumstances obtain in Alphie and Betty's dialogue. On the contrary, it is easy to imagine the discourse occurring under circumstances where both speakers are sincere and

straightforward, and where both speakers have ample evidence regarding the flavors, textures, and aromas of Eggo Waffle Cereal. Given those considerations, the dialogue in (1) starts to look a little bit like the dialogue in (2).

> (2)   (a)   Ivan is tall.
>
>   (b)   Ivan is not tall.

It is easy to imagine a scenario where the speaker of (2a) has in mind the standards of a philosophy department while the speaker of (2b) has in mind the standards of a basketball team. In such a scenario, neither speaker is mistaken, and furthermore, neither speaker need be insincere or sarcastic. As a result, there is widespread consensus among linguists[2] that the word *tall* can pick out different properties—differing in the relevant standard for height—in different contexts. Given that all the same considerations apply to the dialogue in (1), the natural conclusion is that *delicious* can likewise pick out different properties—differing in the relevant standard of taste—in different contexts.[3]

The claim that Alphie and Betty pick out different properties with the same word can explain our intuition that neither is mistaken. But that advantage is undercut by the second intuition. The second intuition, recall, is that Alphie and Betty disagree with each other. What can that tell us about a theory of the word *delicious*? If two speakers really disagree with one another—as it seems that Alphie and Betty do—then there must be something about which they disagree. In the absence of some more complex story, the most natural conclusion is that they disagree over the truth of the proposition asserted by the first speaker. In other words, something like the following principle seems to obtain: if two speakers really disagree with each other, then the propositions they express must be inconsistent. I will refer to that principle, stated a bit more precisely, as *Substantive Disagreement*, or *SD*.

---

[2] See, for example, Kennedy (1999).

[3] The situation is not quite as simple in the case of deliciousness as it is in the case of height. Speakers may situate the threshold for tallness differently, but they will agree on the linear scale along which that threshold varies. By contrast, speakers may situate the threshold for deliciousness differently along a single scale, but they may just as easily employ different scales—possibly incommensurable and not obviously linear—altogether. As I hope will become clear below, these added degrees of freedom are all to the advantage of the view I defend here. Thanks to Ken Walton for emphasizing this point to me.

Substantive Disagreement: If in a discourse, Speaker A and Speaker B really disagree with one another, then there is some proposition *p* such that some utterance by Speaker A expresses *p* and some utterance by Speaker B expresses *q*, where *q* entails *not-p*.[4]

Why should we find SD plausible? Consider the dialogue in (3).

(3) (a) I'm in Madison.

(b) I'm not in Madison.

Suppose that Alphie utters (3a) while in Madison, while Betty responds over the phone by uttering (3b) from her home in Ann Arbor. Like the dialogue in (1), the second expression in this dialogue is the negation of the first. But in dialogue (3) it is clear that there is no disagreement, despite the fact that (3b) is the negation of (3a). The difference is that the expressions in (3a) and (3b) contain contextually sensitive linguistic items: the word *I* picks out different people depending on the speaker. As a result, Betty uses (3b) to deny a different proposition than the one asserted by Alphie via his utterance of (3a). So the proposition expressed by Alphie is *consistent* with the proposition expressed by Betty. Looking to SD, we can observe that despite the similarities between the sentence Alphie utters and the sentence Betty negates, the propositions they express are consistent, and therefore Alphie and Betty do not disagree with one another.

Now consider dialogue (1) again. If Alphie uses *delicious* to express the proposition that Eggo Waffle Cereal has property F, and Betty uses *delicious* to express the proposition that Eggo Waffle Cereal does not have property G—where G is distinct from F—then the contents of their respective utterances are consistent.[5] But if the propositions they expressed were consistent, it would follow by SD that Alphie and Betty do not disagree about Eggo Waffle Cereal. But they *do* disagree. So, Alphie and Betty must pick out the same property with their uses of the word *delicious*.

---

[4] SD requires that Speaker B's utterance *entail that not-p* rather than *express that not-p* because of dialogues like the following:
  (i) Nobody's home.
  (ii) Nuh uh, John is at home.
Such dialogues are clear cases of disagreement despite the fact that (ii) is not a simple negated form of (i).

[5] Of course, certain choices of F and G could still guarantee that the propositions are inconsistent. Those are not the properties we have in mind however. In the cases that are relevant, F will be a property like delicious-for-Alphie and G will be a property like delicious-for-Betty.

4

Arguments very much like the one above have recently been proposed by linguists and philosophers advancing *relativist* semantic theories.[6] According to these linguists and philosophers, *contextualism* about predicates of personal taste—the view that predicates of personal taste pick out different properties on different occasions of utterance—explains the first intuition at the expense of explaining the second. That is to say, contextualists can explain the intuition that neither speaker in a dispute about taste is mistaken. But their explanation of that intuition denies them an explanation of the intuition that speakers in such a dispute are disagreeing.

In response to this dilemma, the relativist proposes, not just an alternative semantics for predicates of personal taste, but a revision to the framework in which such theories are constructed. The details vary, but each relativist proposes some mechanism by which the truth values of sentences are relative, not only to a context of utterance, but also to a judge or context of assessment.[7] Given this radical departure from standard frameworks, it will be worthwhile investigating the motivation for this family of proposals. With that in mind, my focus will be on the motivating argument for relativism and the phenomenon that drives that argument: disagreement.

2. Definition of Disagreement

The principle SD makes reference to situations where speakers disagree with one another, but there are at least two candidates for the relevant set of situations. On the one hand, there are situations where two speakers feel that they are in engaged in some form of dispute. On the other, there are situations where linguistic denial—marked in English by expressions such as *nuh uh*, *nope*, or *no it isn't*—is licensed as a felicitous move in the

---

[6] See, for example, Lasersohn (2005) and MacFarlane (2007). Both will be discussed in more detail below. For similar argumentation, see Stephenson (2007), and, in a related domain, Egan, Hawthorne, and Weatherson (2005).

[7] See von Fintel and Gillies (2008) for a generalized discussion of relativist proposals.

conversation.[8] The two phenomena often go together, but not always. The choice of which phenomenon is relevant to SD will determine whether that principle has any plausibility.

In dialogue (1), the two phenomena go together. It is natural to interpret Alphie and Betty's discourse as a dispute, and furthermore, it is felicitous for Betty to respond to Alphie by denying what he said. But taking oneself to be in conflict with one's interlocutor is not a sufficient condition for the felicity of linguistic denial. Consider the dialogue in (4):

(4)   (a)   I like Dave.

      (b)   I don't like Dave.

For the speaker of (4b), it is infelicitous to assert the content of her utterance in the form of a denial of (4a):

(5)   #Nuh uh, I don't like Dave.[9]

Nevertheless, any number of other indicators of conflict are entirely felicitous.

(6)   Well, I don't like Dave. In fact I think you're dead wrong about him.

Suppose we allow the speakers to make it even clearer that they take themselves to be at odds:

(7)   (a)   I like Dave.

      (b)   Well, I don't like Dave at all. In fact I dislike him intensely. Let's change the subject to something we can agree about.

Even when it is made exceedingly clear that the speakers take themselves to be at odds, linguistic denial remains infelicitous:

_____

[8] As van der Sandt and Maier (2003) observe, denials need not take the form of negative sentences. In response to (i), for example, (ii) is a perfectly effective denial.
      (i)   The cat is not on the mat.
      (ii) Yes he is. The cat *is* on the mat.
Most of my examples of denial will involve negative sentences, but nothing in my argument will depend on that fact.

[9] The expressions *nuh uh* and *nope* are well suited to this discussion because they do not bias the question of which feature of an utterance is being objected to. *No I don't*, *no it isn't*, and similar expressions explicitly target the propositional content of the utterance by negating a shortened version of the same expression, and are therefore less useful. Some English speakers do not have clear intuitions about the felicity of *nuh uh*. Such speakers may find *nope* more familiar, and are welcome to substitute it for *nuh uh* throughout the paper.

(8)   #Nuh uh, I don't like Dave at all. In fact I dislike him intensely. […]

It may be that any situation where denial is felicitous is an occasion where speakers take themselves to be at odds. But cases where speakers take themselves to be at odds, yet cannot felicitously deny one another's assertions are not only possible, but common and natural. In that way, the two phenomena come cleanly apart. Because they are distinct, any discussion of disagreement will need to clarify which phenomenon is intended.

One might be tempted to avoid this ambiguity by allowing SD to retain its present formulation, in which it requires simply that two speakers "really disagree". But that formulation cannot do the work for which SD is intended. If the word "really" is doing any work here, it would seem to suggest that the disagreement is one in which the object of dispute is the truth of the proposition expressed by one of the speakers. In other words, a disagreement between two speakers is *real* only if the expressions uttered by the parties to the disagreement are inconsistent. But the consistency of the expressions uttered is precisely what is at issue. The application of SD in its original form is therefore question begging.

So, given that reference to the notion of "real disagreement" is in this instance question begging, and that disagreement is in any case ambiguous between two sets of phenomena, which of the two phenomena should form the basis of SD? Suppose we choose the intuition of conflict:

Substantive Disagreement*: If in a discourse, Speaker A and Speaker B take themselves to be in conflict, then there is some proposition p such that some utterance by Speaker A expresses p and some utterance by Speaker B expresses q, where q entails *not-p*.

What does the revised principle predict? Consider again dialogue (4), repeated here as (9):

(9)   (a)   I like Dave.

        (b)   I don't like Dave.

Insofar as the speakers of (9a) and (9b) can take themselves to be in conflict—and we've seen that that's pretty far—SD* predicts that the proposition expressed in (9b) must entail

that the proposition expressed in (9a) is false. But that prediction is clearly wrong. The proposition that the speaker of (9a) likes Dave is very much consistent with the proposition that the speaker of (9b) does not like Dave. To be clear, I do not mean to claim that there is no proposition at all forming the basis of this dispute. In the case of dialogue (9), the proposition that *Dave is a nice guy* might very well play just such a role. However, SD* is meant to license inferences about the semantics of the expressions actually uttered in the dispute. If disputes can focus on unexpressed, background propositions—as may be the case in dialogue (9)—then those inferences will fail. The point only applies more strongly if disputes such as those in (9) do not focus on propositions at all—if, for example, they represent some general aversion to differing attitudes. Either way, SD* is false, and any argument that requires it is unsound.

Suppose then that we let SD refer instead to the other phenomenon associated with disagreement—the felicity of linguistic denial.

> **Substantive Disagreement**:** If in a discourse, linguistic denial is a felicitous move for Speaker B in response to an utterance by Speaker A, then there is some proposition *p* such that Speaker A's utterance expresses *p* and Speaker B's utterance expresses *q*, where *q* entails *not-p*.

What kinds of predictions does SD** make?[10] Let's begin by considering some clear cases.

> (10) (a)   The cat is on the mat.
>
> (b)   Nuh uh, the cat is not on the mat.

In dialogue (10), it is easy to imagine that the speakers take themselves to be at odds. More importantly for our purposes, linguistic denial is a perfectly felicitous move for the speaker of (10b). Given that second fact, SD** predicts that the proposition asserted by the speaker of (10a) should be inconsistent with the proposition asserted by the speaker of (10b). That seems to be the right prediction.

---

[10] SD** cannot be quite right as stated because of dialogues like the following.
     (i)   I think it's raining.
     (ii)   Nuh uh, it is not raining.
As (i)-(ii) show, denials can target the embedded clause of a propositional attitude claim. (See von Fintel and Gillies (2008) for a discussion.) Having noted this phenomenon, I will assume that SD**, a principle I hope to critique for other reasons, could be tweaked to account for these data, and will set this issue aside.

What about the other cases we've considered so far? Consider dialogue (3)—repeated here as (11)—and yet again dialogue (4)—repeated here as (12).

    (11)   (a)   I'm in Madison.

               (b)   I'm not in Madison.

    (12)   (a)   I like Dave.

               (b)   I don't like Dave.

The dialogues differ insofar as we can imagine the parties to dialogue (12) taking themselves to be in conflict, while it is harder to imagine the parties to dialogue (11) feeling that way. Nevertheless, the two dialogues have in common that the second speaker cannot felicitously express the content of her utterance in the form of a denial:

    (13)   #Nuh uh, I'm not in Madison.

    (14)   #Nuh uh, I don't like Dave.

Given that denial is not felicitous, SD** makes no prediction about the propositions expressed. That leaves us free in each case to say that—whether or not the speakers are in conflict—the contents of their expressions are consistent. That seems to be the right result. If there's a principle here to be used in the relativists' arguments, SD** seems to be the most plausible candidate.

Because the principle based on intuitions of dispute is false, and because the principle based on felicity of denial has some plausibility, I propose to use the felicity of denial in a stipulative definition of *disagreement*. What I propose is the following definition:

Disagreement =_{df} The relation between speakers that licenses linguistic denial

By defining disagreement in this way, I do not intend to repudiate other uses of *disagreement*. Say, for example, that denials turn out to be licensed in some situations where it seems that speakers are "talking past" one another. There is a well entrenched usage of *disagreement* on which such speakers are not disagreeing with each other. Nevertheless, I will not use the term that way. Even in talkings past, there is a sense in which speakers do disagree with each other. They might describe the situation in retrospect by saying, "We were disagreeing with each other, but it turned out to be

stupid." The advantages of using the term my way will be clarity (achieved by means of stipulation), and the ability to investigate the different *kinds* of disagreement, without fearing the intuitive but vague and potentially question-begging specter of "real disagreement". I will argue in fact that a number of different relations between speakers license linguistic denial. If I'm right, then the conclusion is that there are varieties of disagreement. Some disagreements may be worth engaging in; some may turn out not to be. It will be easy enough to describe the circumstances picked out by other uses of the term; nothing in my arguments will rest on these terminological decisions.

3. Varieties of Disagreement

So the question at hand now is what kinds of relation must obtain between speakers for linguistic denial to be a felicitous conversational move. The most familiar form of disagreement in discourse is exemplified in dialogue (10), repeated here as (15):

(15)   (a)   The cat is on the mat.

          (b)   Nuh uh, the cat is not on the mat.

In such a situation, Speaker A asserts a proposition that Speaker B believes to be false. In response, Speaker B asserts the complement of the proposition expressed by Speaker A. Closely related cases involve pairs of propositions that are inconsistent, even if the second utterance is not a simple negation of the first:

(16)   (a)   No one is home.

          (b)   Nuh uh, John is at home.

(17)   (a)   The box is all red.

          (b)   Nuh uh, it is partly blue.

Disagreements like (15) through (17) have to do with the truth of the propositional *content* asserted in one utterance and denied in the other. For that reason, I will refer to this paradigmatic kind of disagreement as *Content Disagreement*.

Are there other kinds of disagreement besides Content Disagreement? An utterance carries many kinds of information beyond the literal content of the expression

uttered. It has entailments, presuppositions, implicatures. Even the manner in which a proposition is expressed carries information about the speaker and the language being spoken. Each of these modes of information transfer could potentially underlie a dispute between speakers. The dialogues in (16) and (17) show that disagreements over entailments can. But what about the other forms of information that can be carried by an utterance? Can disagreement over less familiar modes of information transfer license linguistic denial? In almost every case, the answer is yes. The following examples should give some sense of the range of denial-licensing disputes.

Presupposition Disagreement:

    (18)   (a)   John remembers going to college.

                (b)   Nuh uh, John never went to college.

Implicature Disagreement

    (19)   (a)   Todd has four kids.

                (b)   Nuh uh, Todd has five kids.

    (20)   (a)   Gary ate some crackers and went to bed.

                (b)   Nuh uh, he went to bed and ate some crackers.

    (21)   (a)   Céline produced sounds closely corresponding to the National Anthem.

                (b)   No she didn't, she *sang* the National Anthem.

Manner Disagreement

    (22)   (a)   Roger won the game 40 to zero.

                (b)   Nuh uh, he won it 40-*love*.

    (23)   (a)   Pedantic jerks correct the [pro-*noun*-ciation] of others.

                (b)   No they don't. They correct the [pro-*nun*-ciation] of others.

In all of these cases, the dialogues not only generate the intuition that the speakers are in conflict, but they also make denial a felicitous conversational move.[11] This despite the fact that none of them of them reflect disputes over the ordinary propositional content of the expressions uttered. Furthermore, in the cases of Implicature and Manner Disagreement, the semantic value of the second expression is not even *inconsistent* with the semantic value of the first. These cases should therefore cast serious doubt on SD**.

In defense of SD**, it might be argued that participants in dialogues like (18) through (23) can be *aware* that their disagreement is about something other than the proposition expressed. In cases of Manner Disagreement in particular, it is at least plausible to think that speakers might know that their conflict does not concern the literal content of their utterances. If non-content disagreements are indeed easily identified as such—as perhaps manner disagreements are—then SD** could still prove useful, provided it is applied with discretion. But the transparency of Manner Disagreement is not shared by all forms of disagreement. Consider two final varieties: *Character Disagreement* and *Context Disagreement*.[12]

Character Disagreement

    (24)   (a)   Burgers come with chips.

              (b)   No they don't. Burgers come with french fries.

    (25)   (a)   I have arthritis of the thigh.

              (b)   No you don't. Arthritis is a joint condition.

    (26)   (a)   Secretariat was a great athlete.

              (b)   No he wasn't. He was a horse.

---

[11] I ignore here the distinction between *descriptive* negation and *metalinguistic* negation, described in Horn (1989), ch. 6. That distinction will play an important role in a complete discussion of the various categories of disagreement. However, it will not affect my arguments here. Metalinguistic negation is non-obligatory in many cases of Presupposition and Implicature Disagreement, including (18), (19), and (20). In cases of Character and Context Disagreement, discussed below, it is not even permitted. If recent work, such as van der Sandt and Maier (2003) is correct, then there is no distinction, and the point is moot.

[12] One note on naming the categories of disagreement. As I discuss below, cases of Character Disagreement involve speakers who disagree about what the character of a word is, or should be. Likewise, cases of Context Disagreement involve speakers disagree about what the linguistically relevant features of the context are, or should be. By contrast, in cases of Content Disagreement, speakers do not disagree about what the content *is*; rather they disagree about whether the content is *true*. This disanalogy aside, I think the labels effectively flag the important features of the categories they name.

In the first dialogue, suppose that the speaker of (a) is British and the speaker of (b) is American. In the second dialogue, suppose that speaker of (a) is from a community that consistently uses the word *arthritis* to refer to sore muscles and let the speaker of (b) be an ordinary English-speaking doctor. Finally, in the third dialogue, let the speaker of (a) be a person who uses the word *athlete* to refer to any animal that engages in sport, and let the speaker of (b) be a person who uses it to refer exclusively to *people* that engage in sport.

In cases of Manner Disagreement, it was clear that the disputes rested on purely linguistic grounds. In cases of Character Disagreement, it can be much harder to draw the line between linguistic and non-linguistic facts. If we look only at dialogues like (24) —a clear case of talking past—we might be tempted to think that these disputes rest purely on misunderstandings about language. But in (25) and (26), the question is much less clear. Certainly, the question of whether arthritis is a muscle or joint condition is an empirical, *non-linguistic* issue. Nevertheless, the question of what a speaker uses the sounds [arthritis] to refer to is an empirical *linguistic* issue: one that in (25) has been settled by stipulation. Yet despite my stipulation, it's hard to imagine the speakers of (25a) and (25b) treating their dispute as a mere talking past. If the linguistic facts were revealed to them, it seems unlikely that the doctor would happily assent to the truth of (25a). If the dispute were purely, unambiguously linguistic—if this talking past were a *mere* talking past—it would be hard to understand the doctor's reluctance. Finally, what are we to say of the speakers in (26)? On the one hand, I've stipulated that they use the word *athlete* differently, making their dispute center on a semantic issue. On the other hand, the only difference between them seems to be whether they think non-humans can be athletes, a question that is not obviously about language.

As these considerations show, in cases of character disagreement, it can be extremely difficult to draw the line between a dispute about non-linguistic facts, and a dispute about how to use one's words. Fortunately for me, I won't have to settle this question either way. For my purposes, three less controversial points will be important: First, cases of Character Disagreement can generate the feeling of conflict in speakers.

The Secretariat case, for example, was originally observed by Peter Ludlow, who heard the argument—and it was an argument—taking place on sports radio.[13] Second, cases of Character Disagreement do license linguistic denial. Whether the dispute turns out to be a mere "talking past" or something deeper, denial is in every case here a perfectly felicitous move in the conversation. Finally, cases of Character Disagreement are potentially very difficult to differentiate from perfectly ordinary cases of Content Disagreement. If, for example, the speaker of (24a) happened to lack the phonological patterns characteristic of British English, one could imagine the dispute carrying on for some time before the speakers realized its true nature. The same point applies all the more strongly to the latter two disputes, which go well beyond mere talkings past. The fact that cases of Character Disagreement are so difficult to distinguish from cases of Content Disagreement is a serious problem for the defender of SD**. If the defense of SD** relies on the accuracy of speaker intuitions regarding the true nature of a dispute, then cases of Character Disagreement show that it cannot succeed.

Context Disagreement

Consider now one last form of disagreement, a category that will be useful in responding directly to the relativist's anti-contextualist arguments. Recall the dialogue in (2), repeated here as (27):

> (27)   (a)   Ivan is tall.
>
>   (b)   Ivan is not tall.

Originally we imagined that the two speakers were in different contexts: one in a philosophy classroom and the other on a basketball court. What we found was that *tall* picked out different standards of height in the two contexts. Because of this property, speakers who already know the salient standard can learn from a use of *tall* about an object's height.. But things can run in the other direction too. Chris Barker describes these so-called *metalinguistic* or *sharpening* uses:

---

[13] Ludlow (2008).

Normally, (1) will be used in order to add to the common ground new information concerning Feynman's height:

> (1)   Feynman is tall.

> But (1) has another mode of use. Imagine that we are at a party. Perhaps Feynman stands before us a short distance away, drinking punch and thinking about dancing; in any case, the exact degree to which Feynman is tall is common knowledge. You ask me what counts as tall in my country. "Well," I say, "around here,.." and I continue by uttering (1). This is not a descriptive use in the usual sense. I have not provided any new information about the world, or at least no new information about Feynman's height. In fact, assuming that *tall* means roughly 'having a maximal degree of height greater than a certain contextually-supplied standard', I haven't even provided you with any new information about the truth conditions of the word *tall*. All I have done is given you guidance concerning what the prevailing relevant standard for tallness happens to be in our community; in particular, that standard must be no greater than Feynman's maximal degree of height. The context update effect of accepting (1) would be to eliminate from further consideration some candidates for the standard of tallness. My purpose in uttering (1) under such circumstances would be nothing more than to communicate something about how to use a certain word appropriately—it would be a metalinguistic use. (Barker 2002, p.1)

Barker goes on to argue that metalinguistic uses of gradable adjectives like *tall* are not only possible, but common. In fact he argues that almost any use of a vague predicate produces some context-sharpening effect, in addition to whatever other information is conveyed.

If context sharpening is a commonly available mode of conveying information, then a natural prediction is that such information is a possible focus of dispute. That prediction turns out to be accurate. Consider (27) again, but let the circumstances be more like those Barker imagines. Gramma has asked Alphie what counts as tall in this area. Alphie responds by uttering (27a), referring to Ivan, who is standing in clear sight of all parties to the conversation. It is easy to imagine Betty disagreeing with Alphie and uttering (27b). In fact, she could quite naturally assert that content in the form of a denial of Alphie's utterance:

> (28)   Nuh uh. Ivan is not tall.

Alphie and Betty's disagreement cannot be about Ivan's height; after all, Ivan is standing right in front of them. What they disagree about is what context they are in. In particular, they disagree about what level of height is the salient standard for tallness.[14] In cases like this, there are two ways to describe Alphie and Bettie's context-based disagreement. On the one hand, we could suppose that there are independent facts about the linguistically relevant features of the context, and that Alphie and Bettie's disagreement is a factual dispute over what the context actually is. On the other hand, we could suppose that the linguistically relevant features of the context are in some ways up for grabs, and that Alphie and Betty offer competing proposals for what the context should be.

This difference is very much worth exploring, but again, for my purposes, three less controversial points will be important. First, like cases of Character Disagreement, cases of Context Disagreement can easily generate speaker intuitions of conflict. And, second, as (28) shows, that conflict can license linguistic denial.[15] Finally, cases of Context Disagreement—like cases of Character Disagreement—differ from ordinary content disagreement in subtle and esoteric ways. Betty, for instance, might happily describe the situation without any indication that anything was out of the ordinary.

> (29)   Alphie and I were disagreeing about whether Ivan was tall. … Yes, of
>   course the dispute was about his height. I told you, we were disagreeing
>   about whether he was tall.

A philosopher or linguist might take issue with Betty's description, but it is hardly obscure why she describes things as she does. The more general point is that speakers cannot be relied upon for accurate judgements about the nature of their disagreements. A defense of SD** that relies on the accuracy of those judgments can therefore not succeed.

---

[14] In the framework of dynamic semantics, Barker gives a formal analysis of how this information is conveyed. However only the existence of sharpening uses will matter for my purposes.

[15] For those keeping track, both Character Disagreement and Context Disagreement seem totally incapable of licensing metalinguistic negation. Although in some ways they reflect disputes about language, they only make available the most robust, proposition-denying (as opposed to *assertibility-denying*) form of negation. Insofar as the distinction is sound, that fact may play some role in explaining speaker blindness about the nature of such disputes.

4. The Basic Anti-Contextualism Argument

        The wide diversity of denial-licensing disputes has important consequences for

the puzzle about predicates of personal taste. As a reminder of how the disagreement-

based anti-contextualism argument goes, consider the following two formulations:

Lasersohn 2005:

> [The context dependent] solution cannot be right […] because of the relation
> between contents and contradiction […]. If I say "Roller coasters are fun," and
> you say "No, roller coasters are not fun," on this analysis, you are not
> contradicting me, because the negated sentence doesn't express the same content
> for you as it does for me. In effect, my utterance means roller coasters are fun for
> me, and your utterance means roller coasters aren't fun for you, and there is no
> conflict or contradiction between those at all—indeed, *there is no reason to think*
> *we disagree in any way, on this analysis*. (Lasersohn 2005, pp. 649. Emphasis
> mine.)

MacFarlane 2007:

> If in saying "apples are delicious" I am saying that they taste good *to me*, while in
> saying "apples are not delicious" you are denying that they taste good *to you*, then
> we are no more disagreeing with each other than we would be if I were to say
> "My name is John" and you were to say "My name is not John".   Intuitively,
> though, it does seem that we are disagreeing. We certainly *take ourselves* to be
> disagreeing. I may say, "Wrong!" or "That's false"—neither of which would be
> appropriate if you had said explicitly that apples taste good to you. (MacFarlane
> 2007, p.18)

Both authors argue that on the contextualist account there is no reason to think that

speakers engaged in a dispute about taste are disagreeing with each other. Their basis for

that claim is the fact that according to the contextualist, there is no single proposition that

is asserted by one speaker and denied by the other. In other words, the basis for their

claim is the principle I've called SD.

        As we saw in Section II, the form of SD that relies on the intuition of conflict—

cases where, as MacFarlane would say, "we certainly *take ourselves* to be disagreeing"—

is massively implausible. As for the version that relies on the felicity of denial, any

dispute from the categories of Implicature Disagreement, Manner Disagreement,

Character Disagreement, or Context Disagreement provides a clear counterexample.

Disagreement from any of those categories license denial, despite the fact that the speakers express mutually consistent propositions.[16]

The upshot of these observations is that neither version of SD is tenable. Any argument premised on SD is therefore unsound. From the fact that speakers disagree—whatever we take disagreement to be—it simply does not follow that the speakers express mutually inconsistent propositions. What the failure of SD gives us first is a burden shifting argument: There is no special reason to doubt that contextualism can handle disagreement data. Disputes that do *not* require inconsistent propositions are common, diverse, and, often enough, not easily distinguishable from disputes that do. If an anti-contextualist wishes to make some specific set of disagreement data problematic, she is obliged to argue that the disagreement at hand requires the expression of inconsistent propositions; that is to say, she is obliged to show that the disagreement at hand falls into the category of Content Disagreement, or possibly Presupposition Disagreement, but not into any of the categories Implicature, Manner, Character, or Context Disagreement. But of course, if she succeeds in doing so, then she has already shown by other means that contextualism is false.

What the failure of SD gives us down the road is the basis for a positive contextualist account of disputes about taste. The first thing to observe about such an account is that it will be disjunctive. Because disagreement data themselves divide into two categories, there are two distinct phenomena to be explained: intuitions of aesthetic conflict, and the felicity of the denial of aesthetic claims. Consider first the intuition of conflict. When Alphie and Betty argue about Eggo Waffle Cereal, they take themselves to be disagreeing. How can contextualism account for that fact? As we've seen the fact that two speakers take themselves to disagree does not require that they express inconsistent propositions. So consider the simplest contextualist analysis of the dialogue in (1), which makes it roughly equivalent to (31):

---

[16] One caveat: depending on how contextual parameters are actually set in discourse, cases of Context Disagreement may or may not provide examples of just this type. In (27a) and (27b), if Alphie and Betty's intentions to use *tall* in a certain way are enough to determine the proposition they've actually expressed, then Alphie and Betty have indeed expressed mutually consistent propositions. If the parameter is set in some more complicated way, that could turn out not to be the case.

(31) (a) Eggo Waffle Cereal tastes good to me.

(b) Eggo Waffle Cereal does not taste good to me.

Relativists like Lasersohn and MacFarlane point out, correctly, that denial would be infelicitous in responses like (32):

(32) #Nuh uh, Eggo Waffle Cereal does not taste good to me.

However, because they do not distinguish between the felicity of denial and the intuition of conflict, they fail to note that such dialogues can nevertheless be part of heated disputes:

(33) Well, Eggo Waffle Cereal does not taste good to me. What's wrong with you? You might as well drink a bottle of Mrs. Butterworth. Can't we agree about anything?

For the contextualist then, the conflicting attitudes that the speakers express is all that is required to explain their "taking themselves to disagree". The fact that in expressing their opposing attitudes they fail to express inconsistent propositions is irrelevant.

What's left for the contextualist to explain is the second phenomenon: the fact that unlike in (31), denial *is* licensed in Alphie and Betty's original dispute. At this point, the analogy between *delicious* and *tall* will again be helpful. In the original *tall* case, we imagined two speakers in two different contexts, with no disagreement. But later, in the discussion of Context Disagreement, it turned out that metalinguistic uses of *tall* are not only possible, but, according to Barker, ubiquitous. Disagreements turning on those uses should likewise be entirely natural, and as we saw, constructing such disputes involves an entirely straightforward extension of Barker's original example. It seems natural then to extend the metalinguistic analysis of *tall* disputes to disputes about taste. A detailed Context-based account of aesthetic disputes is beyond the scope of this paper largely because of the issue I raised earlier—the question of whether context disputes are about what the context is, or rather what the context should become. That is an issue I prefer to remain neutral on for the time being, largely because I think that the proper theory will involve a bit of both, and saying just what the distribution should be is a paper in itself. Nevertheless, it will be possible to set that issue aside and still give a rough sketch of what any Context-Disagreement based account would look like.

As Lasersohn himself points out, the contextualist need not tie predicates of personal taste to the speaker exclusively. The salient standard might be the standard of some group, though that group will under most circumstances at least contain the speaker and other parties to the conversation.[17] Suppose then that Alphie utters (1a) with a contextually salient group standard in mind. The standard will be sensitive to the tastes of individuals other than Alphie; but Alphie's tastes will still factor in. What he expresses then, is the proposition that, according to the standards of that group, Eggo Waffle Cereal is delicious. Now Betty, whose personal standard of deliciousness is not satisfied by Eggo Waffle Cereal, similarly expects her own standards to be included in the salient group standard. She therefore has very good reason to reject Alphie's utterance; after all, she knows that it is false that Eggo Waffle Cereal satisfies the group standard determined by the intersection of Alphie's and her own tastes. In the original case then, the contextualist does not even need to appeal to metalinguistic usage.[18] Given each speaker's natural assumption that the group standard will involve themselves and other parties to the conversation, Alphie's utterance is false and Bettie knows it.

But what if the order is reversed? Suppose Betty utters (34a) and Alphie responds with (34b):

(34) (a)   Eggo Waffle Cereal is not delicious.

(b)   Oh yes it is!

As Lasersohn points out[19], a similar story will not apply in this case, since Alphie will know—in virtue of Betty having said what she does—that any group standard involving both parties to the conversation will not be satisfied by Eggo Waffle Cereal. But the contextualist is not out of options at this point. In addition to its descriptive effects, Alphie's utterance has metalinguistic effects regarding the appropriate usage of *delicious*. In particular, his utterance conveys the information that in the present context, the

_____

[17] It turns out that containing the speaker is not a strict requirement. That said, such cases present no problem for the contextualist, so I will set them aside. See DeRose (1991) for discussion.

[18] Though the contextualist *can* appeal to metalinguistic usage even in these simple cases. It may be that in practice disputes can go either way.

[19] Lasersohn 2005, p. 651.

relevant standard is (or at least should be) such that Eggo Waffle Cereal does satisfy it. This account not only makes sense of the fact that Alphie and Betty disagree, but it also suggests—accurately, I think—that Alphie in (34b) is being more belligerent than Betty in dialogue (1). In (1), Betty points out that Alphie is mistaken in his claim about how the group standard rates Eggo Waffle Cereal. In (34) by contrast, Alphie intentionally excludes Betty from that group. I think this may explain why, in the original version of this case, Lasersohn felt it necessary to include an exclamation mark. Alphie's response is considerably less natural without it. Whatever the intensity of the dispute in (34), the dispute belongs, along with metalinguistic disputes about tallness, in the category of Context Disagreement. And as was demonstrated in Section III, disputes belonging to that category are capable not only of *feeling* like disputes, but also of licensing linguistic denial.

5. Advanced Anti-Contextualism Arguments

For the contextualist then, ordinary disputes about taste present no problem at all. Some can be described as simple cases of Content Disagreement, while others can be described with reference to Context Disagreement, a form of dispute the existence of which is well motivated by data from philosophically more innocent domains. However, both Lasersohn and MacFarlane attempt to cut off the later kind of explanation by turning to another kind of data. Although their arguments are superficially quite different, both try to make trouble for Context Disagreement analyses by considering disputes at some remove from an actual conversation.

I'll consider MacFarlane's argument first. MacFarlane is interested cases such as the following:

> When I was ten, I used to go around saying "fish sticks are delicious" (and meaning it!). Now I say "fish sticks are not delicious." It seems to me that I disagree with my past self. But I am not involved in a *conversation* with my past self. (MacFarlane 2004, p. 20)

The difference between MacFarlane's fish stick case and the cases of disagreement we've considered so far is of course that in the fish stick case, there is no conversation—and thus no conversational context to fight about. MacFarlane's point is that context disagreement explanations of disputes about taste require a conversational context. That means they cannot apply in cases of disagreement where there *is no* conversation.

Recall however that MacFarlane never distinguishes between the two kinds of disagreement data: the intuition of conflict, and the felicity of denial. If we consider the two kinds of disagreement phenomena separately, it becomes clear that MacFarlane's case presents no problem for the contextualist. Recall that in Section IV, disagreement about context was called in to explain, not the intuition that Bettie and Alphie disagreed with each other, but rather the fact that in their conversation, linguistic denial was a felicitous move. Note however that the licensing of denial is a moot point in MacFarlane's case, for the very reason that he raises it: there is no conversation. With no conversation occurring between the parties to the disagreement, there is no denial licensing to be explained, and therefore no need for the context disagreement solution.

That leaves to be explained the other kind of disagreement phenomenon: the intuition of conflict. MacFarlane is right that there is some kind of conflict between his present self, and his previous, fish-stick-liking self. But as we've seen, that kind of intuition does not require that inconsistent propositions be expressed. Again, the simplest version of contextualism predicts that young MacFarlane and old MacFarlane's utterances are roughly equivalent to (35) and (36).

(35)   Fish sticks taste good to me.

(36)   Fish sticks don't taste good to me.

MacFarlane is right that no context disagreement story could apply here, since the utterance of (35) takes place many years before the utterance of (36). But the context disagreement story was never required to explain the intuition of conflict.

(37)   Fish sticks don't taste good to me. I can't believe I used to like them. Now that I've experienced real seafood, I can see what a rube I was!

As (37) shows, and as (4), (6), (7), (31), and (33) showed before it, the intuition of conflict carries no requirement that speakers express inconsistent propositions. As long

as speakers express conflicting attitudes, the consistency of the propositions expressed is irrelevant.[20] The contextualist therefore has all the resources she requires to explain MacFarlane's cases. Where she requires the context disagreement explanation, it applies. Where it does not apply, she does not require it.

Lasersohn's argument musters a different kind of case, but operates on the same principle as MacFarlane's. Lasersohn is interested in reports of disagreement, like that in (38).

> (38)   John thinks that roller coasters are fun, but Mary thinks that roller coasters are not fun.

Lasersohn points out that (38) sounds very much like a report of a disagreement. But, he claims, the contextualist cannot predict that reading.

To support his claim, Lasersohn gives two arguments. The first argument relies on a well known bit of semantic speculation: Kaplan's ban on monsters, i.e. semantic operators that shift the context. Lasersohn argues that if predicates of personal taste make indexical reference to a contextually salient standard, and if logical operators that shift the context are impossible, then the contextualist cannot make the correct prediction about (38). I'd like to avoid getting into a long discussion of the ban on monsters. I will note however that the ban has lately come into question from a number of sources[21], and that even if we take it for granted, the ban on monsters cannot possibly do the work Lasersohn wants it to do here. Lasersohn argues that given the ban on monsters, the contextualist must predict that (38) has the following two readings exclusively:

> (39)   John thinks that roller coasters are fun-for-John, but Mary thinks that roller coasters are not fun-for-John.

> (40)   John thinks that roller coasters are fun-for-Mary, but Mary thinks that roller coasters are not fun-for-Mary.

---

[20] *Why* situations like this generate the intuition of conflict is a question that I don't have the answer to. See Egan (forthcoming) for a helpful discussion of the issue. Fortunately, for my purposes here it makes no difference why we have those intuitions. It is enough to observe that we do have them across a wide range of cases where speakers express mutually consistent propositions.

[21] See, for example, Schlenker (2003) and Anand and Nevins (2004).

If those really were the predictions the contextualist were forced to make, this would be bad for the contextualist. But they are not. The first thing to note is that if we are really taking the ban on monsters seriously, we will not merely disallow shifting the contextually salient standard *between* the two clauses; we must disallow shifting the standard in any way at all. The only available reading should thus be (41):

(41)　John thinks that roller coasters are fun-for-the-speaker-of-(41), but Mary thinks that roller coasters are not fun-for-the-speaker-of-(41).[22]

If the ban on monsters applied to contextually variable adjectives in the same way it applies to normal indexicals, then (41) would be the only reading the contextualist could predict for (38). But the ban, if it is to be at all plausible, cannot possibly apply in that way. Again, the comparison with the gradable adjective *tall* will be helpful.

(42)　Sneezy the Dwarf thinks that Ivan is tall, but Michael Jordan thinks that Ivan is not tall.

There is a natural reading of (42) on which the beliefs being reported are Sneezy's belief about whether Ivan is tall by Sneezy's standards, and Michael Jordan's belief about whether Ivan is tall by Michael Jordan's standards. The contextually variable semantics of gradable adjectives like *tall* are well established. The mistake here is to think that Kaplan's ban on monsters prevents any contextually variable term from exhibiting any sensitivity to local linguistic context.[23] Words like *tall* shift easily to take on the standards that are made relevant by local features of the expression. There is therefore no reason at all to think that if words like *fun* are similarly contextually sensitive, they could not do the same thing. That leaves the contextualist free to predict a more natural reading for (38), one on which John's belief is about funness-for-John, and Mary's belief is about funness-for-Mary.

　　Having stripped away these extraneous bits of Kaplania—and Lasersohn himself acknowledges that they not universally accepted—the remaining argument is strikingly similar to the arguments we've encountered already.

---

[22] Thanks to Thony Gillies for emphasizing this point to me.

[23] For an even simpler example of such sensitivity, consider "John put the rough lens on the smooth table." Clearly, standards for gradable adjectives must be able to shift over the course of a sentence. (Thanks to Andy Egan for this example.)

> Of course there are non-Kaplanian semantic theories which allow operators to shift the context.... But simply allowing the context to shift during the sentence will not solve this problem, as long as these context shifts alter the contents of the subordinate clauses. If we interpret the first subordinate clause relative to John, and the second one relative to Mary, then [(38)] will not report John and Mary's thoughts as contradicting each other, contrary to intuition. (Lasersohn 2005, p. 661)

Because the contextualist proposes that the uses of *delicious* in the two clauses are relative to different sets of standards, there is no way for her, according to Lasersohn, to predict our intuition that John and Mary's thoughts conflict with each other. This is the same SD-based argument we encountered earlier, and as we've seen, it is unsound, whatever form of SD is intended.

6. The Puzzle, Redux

Supposing that something like the contextualist story is right, what can we now say about the more general puzzle? One thing we've made sense of is how two speakers who each say something true can nevertheless take themselves to disagree. Making sense of that possibility goes a long way towards explaining how aesthetic disputes can sometimes seem so defective, and other times seem entirely legitimate.
Even when each speaker utters a true proposition, their utterances can carry very different messages about the aesthetic standards being applied. In such a situation, neither speaker utters a falsehood, and in that sense, neither is mistaken. Nevertheless, there is a real point of contention between them.

But those conflicting messages also point to an aspect of the aesthetic puzzle that remains, emphatically, open. When two speakers disagree over what the relevant aesthetic standards are, or what the relevant aesthetic standards should be, what is there to adjudicate their dispute? The fact that on this picture those suggestions are transmitted metalinguistically does not make the question of adjudication any easier. To make the point vivid, suppose, with a nod to Hume, that Alphie utters (43a) and Betty replies with (43b):

(43)   (a)   The poetry of John Ogilby is even more beautiful than the poetry of John Milton.

       (b)   No it isn't. The poetry of John Ogilby is not as beautiful as the poetry of John Milton.

By positing indexical reference to a range of aesthetic standards, the contextualist can offer a range of explanations for the disagreement in (43). And given the many directions that our aesthetic disputes can take in real life, it is all to the advantage of the contextualist that there are a variety of accounts available. What those accounts don't provide, however, are the resources to decide who is right. Alphie's claim, after all, is meant to be crazy. To be sure, the contextualist is not committed to claiming that Alphie's utterance is sane. After all, there are more ways to be wrong than to say something false. Alphie's claim might be true, but true in virtue of referring to an outrageous set of standards. The question of what it is that makes those standards outrageous, however, remains unanswered.

Chapter 2

Disagreement, Denial, and the Modes of Communication

When conversational partners find themselves at odds, they can differ with respect to, among many other things, their beliefs about the proper understanding of the words they use, and their beliefs about linguistically relevant features of the conversational context. I call these types of conflict *Character Disagreement* and *Context Disagreement*, respectively. With reference to cases discussed in recent work by Peter Ludlow and Chris Barker, I attempt to locate these forms of disagreement within a broader understanding of the ways in which information is communicated, arguing that utterances routinely convey information about the proper deployment of linguistic items, a usage I refer to, following Barker, as *sharpening*. I observe that, contrary to what might be expected, denials licensed by disagreements over the proper deployment of linguistic items show none of the markers of so-called "metalinguistic" negation. This fact can be understood as parallel to a generalization originally made in Horn (1988) regarding denials of Q-based and R-based implicatures. That such a pattern generalizes to Character and Context Disagreements suggests, along with other considerations, that sharpening usages are themselves a form of implicature.

1.   Dimensions of Information Transfer

Suppose that, in April 2009, Alphie utters sentence (1) in a conversation with his boss, Bobby. Let's assume further that Alphie's utterance is in crisp Standard American English of the kind that might be heard on a national news broadcast.

(1)   Recently, sir, eleven Democrat members of congress voted against the stimulus package.

Based on the evidence made available by Alphie's utterance in this context, along with some background knowledge and the assumption that Alphie is well informed and a good source of political wisdom, Bobby can infer a rather vast array of conclusions[24]. For a few examples among many, consider (i)-(x).

(i)   Eleven Democratic members of congress voted against the American Recovery and Reinvestment Act of 2009 (ARRA), and that vote happened not too long ago.
(ii)   Somebody voted against ARRA.
(iii)   There were at the time of the vote at least eleven Democratic members of congress.
(iv)   It is not the case that twelve Democratic members of congress voted against ARRA.
(v)   The Democratic Party is somehow sinister or untrustworthy.[25]
(vi)   The phrase *stimulus package* refers, in this context, to the economic stimulus bill that was voted on in the United States House of Representatives in February of 2009, and not, for example, to the stimulus package approved by the Chinese State Council in 2008.
(vii)   The contextually determined threshold for an event's being *recent* is at least two months or so before the time of the utterance.
(viii) Bobby stands in a position of authority with respect to Alphie.
(ix)   Alphie is not Australian.
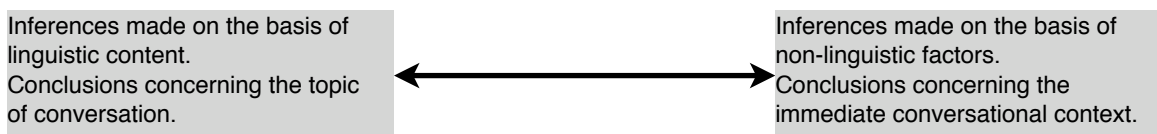(x)   Alphie does not have bronchitis.

These inferences range from conclusions reached almost entirely on the basis of the literal, linguistically encoded content of Alphie's utterance to information that has nothing at all to do with that content, with a great deal of information falling somewhere in between. Likewise, the conclusions themselves vary in content from propositions that concern, exclusively, Congress and the stimulus package, to propositions that concern, exclusively, Alphie, Bobby, and their immediate conversational context. These distinctions are imperfect: I do not have anything like a precise measure of "degree of

---

[24] Throughout the paper, I refer to *inferences that can be made* and *conclusions that can be drawn* on the basis of an utterance. Sometimes a similar observation is couched in terms of the range of information that a *speaker can communicate* via an utterance. That way of putting things is clear enough, but on certain readings it can artificially restrict the range of information to be considered. The intuitions here are not robust, but we might not tend to say that a speaker *communicates* to her interlocutor that she does not have bronchitis. Nevertheless, that she does not have bronchitis is very much something that her interlocutor can learn from her utterance.

[25] Political attitudes suggested in this or other examples do not necessarily reflect the views of the author. Especially views suggested by the annoying phrase *Democrat member of congress*.

reliance on linguistic content"; nor, of course, does anything prevent Alphie from subverting the second distinction by uttering expressions the literal content of which concerns himself, Bobby, and their immediate conversational context. Nevertheless, it is useful to impose some kind of structure on the space of channels along which information can be transferred by an utterance. It will not be important to me to make fine grained distinctions, and so I will present the ordering in what strikes me as a plausible way. Other ways of imposing structure on the same space may be useful for other purposes, and if someone finds a good way to measure reliance on linguistic content and on that basis orders things in a slightly different way, I'm happy enough to revise my list.

So let's allow that, imposing structure on the vast number of things we can learn from an utterance, there is something like a spectrum.

| Inferences made on the basis of linguistic content. Conclusions concerning the topic of conversation. | ⟷ | Inferences made on the basis of non-linguistic factors. Conclusions concerning the immediate conversational context. |

At the left end of the spectrum are inferences made on the basis of linguistic content. Typically, those inferences will yield information pertaining to the topic of conversation, whatever that may be. At the right end of the spectrum lie inferences that are made on the basis of non-linguistic factors, which typically yield information pertaining to the immediate conversational context.[26] The many inferences that fall in between rely on some mix of linguistic and non-linguistic evidence, and concern topics that concern a mix of the events under description and the events characterizing the speech situation. Understandably, semanticists and pragmaticists have focused on the left side of the spectrum. For example, the proposition in (i) is, simply, the semantic content of Alphie's utterance. This is sometimes thought of as *what is said* by Alphie[27], though this way of

---

[26] I am hardly the first to draw these distinctions. They go back, at least, to Peirce's distinction between *index* and *symbol* and, more recently, to Jakobson's work distinguishing between the *conotive* and the *emotive* functions of language—a spectrum at the center of which, tellingly, falls the *metalingual* function. See Jakobson (1956). For a more modern and vastly more precise formulation of similar distinctions, see Potts (2005)'s discussion of the notion of *at issue content*, or Roberts (1996) on the notion of the *question under discussion*. See also Potts (2007) for a related discussion of the "expressive" dimension of language use.

[27] See Stanley (2002) for discussion.

putting things is a bit misleading since Bobby might reasonably report Alphie as having said any number of things near the left side of the spectrum. Moving towards the center, the proposition in (ii), while it is logically weaker than the proposition in (i), requires more work from Bobby, who must draw a logical inference on the basis of (i). The inference of (iii)—one of the propositions that Alphie presupposes in making his utterance—goes a bit farther, involving considerations of common ground, the possibility of accommodation, and similar notions. Finally, to infer the proposition in (iv), Bobby relies not just on the literal meaning of Alphie's utterance, or on its entailments, or on its presuppositions, but also on a complex, if by now familiar, set of considerations of conversational etiquette and felicity.

More remote from content strictly construed, but still very much meaning-implicating, is Bobby's conclusion about the Democratic party. To infer the proposition in (v), Bobby must be sensitive to fairly subtle connotative facts about the words Alphie choses. Why use the loaded form *Democrat member of congress*, rather than the more widely accepted *Democratic member of congress*? The use of *Democrat* as an adjective is limited almost entirely to opponents of Democrats and Democratic policies. Democrats themselves tend to resent it. If Alphie is politically savvy and conversationally cooperative, then he must have some good reason for choosing a manner of speaking that signals solidarity with their opponents. If his political judgements are reliable, then there must indeed be something distasteful about Democrats. The connotative facts licensing this kind of reasoning may or may not be encoded in the lexical entries for the words Alphie choses, but they're nevertheless part and parcel of their *meaning*, broadly construed. The information that is inferred, while it is deduced by way of assumptions about Alphie and his trustworthiness and, plausibly, does not affect the truth conditions of the expression, remains very much focused on the events and actors under description, and not merely on the immediate conversational context.

Consider now the other end of the spectrum. At the right end of the spectrum, both the resources required to make inferences and the content of those inferences have changed. To deduce that Alphie does not have bronchitis, Bobby need not pay any

attention at all to *what Alphie says,* in the sense described above. On the contrary, he requires only his background knowledge of bronchitis and the human vocal tract, together with his knowledge of the fact that Alphie successfully vocalized. As for the conclusion that Bobby reaches on this basis, the fact that Alphie does not have bronchitis has nothing at all to do with Democrats or the stimulus package. In order to reach the conclusion in (ix)—the fact that Alphie is not Australian—Bobby need not know much (or, indeed, anything at all) about what Alphie has said. But he does need to know a bit about the speech patterns of the typical American English speaker, and how those patterns compare to other dialects. The linguistic basis for such a deduction is thin, in the sense that it draws on only certain aspects of Bobby's linguistic knowledge, but ineliminable. Bobby need not know anything about the *meanings* of Alphie's words; Alphie could have been reciting "Jabberwocky", or reading a list of words that he'd just invented. But the information that licenses the inference of (viii) is nevertheless *linguistic*—it's simply phonological rather than semantic. So long as Bobby can assume that Alphie was speaking, rather than, say, imitating animal noises, he will have enough information to observe the phonological features and processes characteristic of a fluent speaker of American English.

Bobby's inference about the power dynamic between himself and Alphie requires even more in the way of linguistic knowledge. Alphie's use of *sir* doesn't alter the truth conditions of the expression he utters. But *sir* is a word with a meaning and it is in virtue of its having the meaning it does that Bobby is able to reach the conclusions about Alphie's show of respect. Along the same lines are brands of inferences that are available in many languages, but not typically in English. Utterances in languages that have sophisticated systems of honorifics, for example, will allow for complex inferences about the speaker and her perceived relationship to other conversational participants.[28] The Papua New Guinean language Hua marks verbs grammatically to indicate the relative elevation of speaker and hearer, allowing a listener to discern whether she is uphill or

---

[28] See, for example, Inoue (1979)

downhill from the speaker without any of the usual types of evidence.[29]   Like connotative

facts, honorifics and similar markings typically have little to do with the proposition

expressed by an utterance, yet are in a non-trivial sense part of the meaning of the

sentence uttered. It is clear that we have worked our way some distance back towards the

linguistic end of the spectrum.

2.   The Center of the Spectrum

        Generally speaking, the information in (i)-(v) is about the set of facts that Alphie

aims to describe with his utterance. And generally speaking, the inferences in (viii)-(x)

are about the set of facts comprising the speech situation in which Alphie's utterance is

embedded. In the middle, between the two extremes, are (vi) and (vii), each of which

represent a mixing of the two, and thus represent for Bobby an act of solving for both sets

of facts. If the facts about the speech situation are one way—if Alphie means certain

things by his words, and certain linguistically relevant features of the context are such

and such—then he describes the facts about the vote in one way. If the facts about the

vote are taken to be different, then so too must be the facts about Alphie's use of his

words, and the linguistically relevant features of the context in which he speaks. If his

phrase *stimulus package*, for example, refers to the package passed by the Chinese

government, then the proposition he expresses is false. If by *recently,* he means a span of

time going back exactly two weeks, then the proposition he expresses is false. Knowing

those facts, Bobby can reach the conclusions described in (v) and (vi), provided he is

willing to interpret Alphie's words charitably. Recent work by Peter Ludlow and Chris

Barker brings to the fore the importance of this kind of information transfer. I'll briefly

describe the cases they consider in order to give a more fleshed out sense of what's going

on in (vi) and (vii) above.

        In "Cheap Contextualism", Peter Ludlow argues that the "common coin" features

of word meaning—the parts of word meaning that are constant from conversation to

---

[29] Haiman, (1979)

conversation—are thin. By "thin", he means that the meaning of any given word is highly underspecified by the linguistically encoded, stable features of the lexical entry.[30] The unspecified aspects of the meaning are, according to Ludlow, fleshed out on the fly by the participants in a conversation. This general thesis, which he calls the *dynamic lexicon*, will not be my focus here, but I will consider some of the cases Ludlow mobilizes in arguing for it.[31] Ludlow considers first the case of two sports fans debating the question of whether Mickey Mantle or Barry Bonds is a better baseball player.

> Well, one of them hit more home runs, but the other was on more championship teams. One of them may have cheated by using steroids. Should that be a factor? What is really up for grabs here is the question of what counts as a "good" baseball player - it is about the meaning of 'good'. (p. 9)

Ludlow considers a similar case to argue that more is at stake than the precisification of vague predicates.

> [C]onsider the following case from a dispute I heard on WFAN (a sports talk radio station in New York) when Sports Illustrated announced its "50 greatest athletes of the 20th Century". Some listeners called in complaining that a horse - Secretariat - had made the list, while host Mike Francessa defended the choice. Clearly this is a dispute about what should be in the extension of 'athlete', and the callers wanted to argue that a horse had no place here. It is not as though the dispute would be resolved if Secretariat were a little bit faster or could throw a baseball, so it seems hard to imagine that these are vagueness cases. (p.10)

Ludlow's choice of an argument about *goodness*—even if it is athletic and not moral goodness—brings with it the philosophical baggage of metaethics. Ludlow is happy to take that on, but for now, I will focus on the less contentious question of what it takes to qualify as an athlete.

In both cases, Ludlow considers sets of speakers who take themselves to disagree, a situation I will return to in detail below. But consider first the simpler case in which the conversation proceeds amicably. Suppose that Alphie utters sentence (2).

---

[30] Ludlow's is one theory among many that posit underspecification of meaning to one degree or another. For some other examples, see von Fintel and Gillies (forthcoming), and the works cited there.

[31] I take the arguments in this paper to be entirely consistent with the dynamic lexicon, and in fact to provide some degree of support for it. For now, I'll remain neutral however. My arguments go through even if one considers cases of the kind Ludlow describes to be unexotic matters of precisification, nominal restriction, or polysenym choice.

(2)   Secretariat was an athlete.

Bobby ordinarily uses *athlete* to refer, exclusively, to people that engage in athletics, and assumes that others use it similarly. However, he also assumes that Alphie is a fairly sensible person. Hearing Alphie's utterance, Bobby *could* choose to hold his semantic assumptions fixed. If he does so, he will conclude that Alphie has said something false, indeed something a bit crazy.

Alternatively, Bobby could choose to hold his assumptions about Alphie fixed. If he does that, then he will conclude that Alphie uses the word *athlete* differently than he does. Bobby's reasoning could proceed in a number of ways, but let's suppose for now that it works something like this:[32] Bobby's default assumption is that his interlocutors use their words in the same way that he does. Call this the *presumption of linguistic similarity.* In the case of Alphie's utterance of the expression in (2), his presumption would lead Bobby to conclude that Alphie has said something grossly defective. That conclusion might be right, but it is dubious enough to be worth weighing against the presumption of linguistic similarity. Calling that presumption into question, Bobby must consider what Alphie might have meant by *athlete.* One way of doing this is to formulate a set of candidate extensions for Alphie's word *athlete.* Extensions that differ massively from Bobby's own are unlikely to be the right extension to ascribe to Alphie's usage. But perhaps there is some candidate extension E such that E is not massively dissimilar to the extension of Bobby's own word *athlete,* and yet if E were the extension of Alphie's word *athlete,* then what Alphie said wouldn't be so crazy. Bobby realizes that there is such a candidate extension: the set, roughly, containing animals (not just people) that engage in sport. If *that's* what Alphie meant by athlete, then what he said wouldn't be crazy; it might even be true. If Bobby is willing to go so far as to defer to Alphie's knowledge of sports and the preferred ways of speaking about sports, then he may even infer that *athlete* is best used in the way that Bobby uses it.

In reasoning along the lines I've described, Bobby draws heavily on both linguistic and non-linguistic evidence. He must rely on assumptions about Alphie (that

---

[32] The story I tell here diverges from how Ludlow would describe matters on the dynamic lexicon theory. Either story will be entirely compatible with my arguments regarding denial and negation.

he's probably not crazy, that he may be worthy of deference), about the meanings of Alphie's words (what should be considered candidate extensions of *athlete*), and even about Secretariat. (If Alphie had said instead that Garfield the lazy cat was a great athlete, it would make an important difference to Bobby's reasoning.) Assuming that he concludes that Alphie uses the word *athlete* differently than he does, Bobby will have to decide whether to defer and adopt that usage or to put up a fight. I'll consider that decision and the consequences of putting up a fight in later sections.

Consider now a different kind of case, this one described in Barker (2002). In that paper, Barker describes what he calls *metalinguistic* or *sharpening* uses of scalar predicates. He introduces the phenomenon with an example involving the predicate *tall*:

> Normally, [3] will be used in order to add to the common ground new information concerning Feynman's height:
>> [3] Feynman is tall.
> But [3] has another mode of use. Imagine that we are at a party. Perhaps Feynman stands before us a short distance away, drinking punch and thinking about dancing; in any case, the exact degree to which Feynman is tall is common knowledge. You ask me what counts as tall in my country. "Well," I say, "around here, ..." and I continue by uttering [3]. This is not a descriptive use in the usual sense. I have not provided any new information about the world, or at least no new information about Feynman's height. In fact, assuming that tall means roughly 'having a maximal degree of height greater than a certain contextually-supplied standard', I haven't even provided you with any new information about the truth conditions of the word tall. All I have done is given you guidance concerning what the prevailing relevant standard for tallness happens to be in our community; in particular, that standard must be no greater than Feynman's maximal degree of height. (p. 2)

Unlike Ludlow, Barker does not consider cases where two speakers disagree about the linguistic information communicated by the utterance. We will consider such cases in detail below. At this point, however, it is enough to observe that there is a close parallel between Ludlow's case and Barker's. Just as a listener can infer facts about how a speaker uses her terms, so too can a listener infer facts about the speaker's beliefs about linguistically relevant features of the context. And just as the listener may go on to defer to the speaker in the extensions she assigns to the relevant terminology, so too might she

defer to the speaker in the way that she fills in the contextual parameters of context-sensitive terminology.

These modes of information transfer fall directly at the center of the spectrum on both strategies for ordering the space. The inferences are made on the basis of both linguistic and non-linguistic evidence. And the conclusions they license concern just those facts about the immediate conversational context that serve to determine what was said about the events under description. Barker describes his cases as *metalinguistic* or *sharpening uses of scalar predicates,* leaving room for the idea that there may be metalinguistic or sharpening uses of other kinds of linguistic items as well. I'll take him up on that, and suggest that there are *metalinguistic* or *sharpening* uses of ordinary predicates—predicates like *athlete*, for example. With that in mind, I will use both *metalinguistic usage* and *sharpening* to refer to any situation where, from an utterance that may have been *about* something else altogether, a listener can draw inferences regarding the linguistic facts of the immediate conversational context.[33] Where it is important to distinguish between Barker-style cases and Ludlow-style cases, I will distinguish *context-sharpening* and *character-sharpening*[34] instances of metalinguistic usage. Just what impact the special features of sharpening has on the discourses in which it appears will be discussed below.

3.   Disagreement Across the Spectrum

We've seen that a rather vast array of conclusions can be reached by a listener on the basis of a speaker's utterance. And we've seen that some of those inferences depend heavily on the linguistically encoded content of the utterance, while some of them do not

---

[33] This usage of *metalinguistic* is not to be confused with other uses in philosophy of language—for example, discussions of Stalnaker's *metalinguistic* or *metasemantic* understanding of semantic matrices. See Stalnaker (1978) and Braddon-Mitchell (2005).

[34] On the dynamic lexicon, what is thought of as ordinary contextual variability is explained in just the same way as subtle variations in Kaplanian character. Both are explained as products of the on the fly process of meaning determination. If it turns out that there is no distinction to be made between character- and context-sharpening, that's fine with me. But a very precise keeping-track of the distinctions among polysemy-resolution, precisification, nominal restriction, and contextual parameter setting is also entirely consistent with my goals in this paper.

depend on that content at all. And we've seen that some of them require reasoning that draws heavily on both kinds of considerations. One fact that is true across the entire spectrum is that the listener, having realized that a certain inference can be drawn from the utterance, may not find the information plausible; she may in fact disagree. The term *disagreement* itself plays a large enough and varied enough role in philosophy that it will be worth saying just a bit about what it means. Often, *disagreement* is used in such a way that disputes over information other than the proposition expressed do not qualify. It is that usage which is intended in claims such as that in (4).

> (4)   Alphie and Bobby were arguing over whether to go to the bank, but they
>   meant different things by *bank*, so they were not really disagreeing.

The usage of *disagree* in (4) is sensible enough, but it is not how I will use the term. There is an equally sensible usage according to which Alphie and Bobby *do* disagree:

> (5)   Alphie and Bobby disagreed about whether to go to the bank, but the
>   disagreement turned out to be pointless since they meant different things
>   by *bank*.

Because of its disparate usages, it will be important to be precise about the meaning of *disagree* but exceedingly difficult to adhere perfectly to any common usage. For my purposes, *disagreement* will describe any situation at all where speakers take themselves to be at odds over some bit of information communicated via an utterance. That will include situations, like Alphie and Bobby's in (4) and (5), in which the disagreement is based on a misunderstanding. In using the term this way, I do not intend to repudiate the other senses in which it used, nor do I intend to offer my own usage as a substantive thesis on the proper analysis of the concept *disagreement*. Having made those qualifications, I think my way of using the term has as good a claim on intuitive, common

as many alternatives[35], and that furthermore it will allow for a wider-ranging discussion of the ways in which speakers can find themselves at odds.

With this notion of disagreement in mind, it becomes trivial that any kind of information that can be inferred on the basis of an utterance can form the basis of a disagreement between speakers. That's just to say that a listener may have reason to object to almost anything they may recognize as a possible inference. Just as a listener can take issue with a proposition expressed that she deems to be false, so too can she dispute a false implicature, or a misleading connotation. Even information channels that have little or nothing to do with the proposition expressed can be co-opted to mislead, or can relate information with which a listener may take issue. A speaker might use the phrase *Democrat senators* to give you the impression that she is a Republican, though she is not. She may adopt an outrageous Swedish accent to convince you that she is not a native English speaker, though she is. She could employ an elaborate system of recordings and ventriloquism to speak as if she did not have bronchitis, though she does. Even the fact that a speaker speaks a particular language can be falsely communicated, as the following dialogue demonstrates.

> Brian: Hola, um, me llamo es Brian. Uh, let's see. Nosotros queremos ir con ustedes...
> Man: That was pretty good. But actually when you said "me llamo es Brian" you don't need the "es". Just "me llamo Brian".
> Brian: Oh, you speak English.
> Man: No, just that first speech and this one explaining it.
> Brian: You're kidding, right?
> Man: Que?[36]

---

[35] There seems to be a connection between intuitions about *disagreement* and intuitions about *communication*. Insofar as we can make sense of the idea that Alphie *communicates* the fact that he does not have bronchitis, it seems that we can make sense of Bobby *disagreeing* in something like the ordinary sense. Since the question of which items on the spectrum are *communicated* and which are *merely inferred* by the listener is itself rather arbitrary, so too is the question of which kinds of objections count as *disagreement* in the ordinary sense. I take that to be strong motivation for using *disagreement* in a technical sense, where, by stipulation, it can apply across the spectrum. For an extended discussion of speaker intuitions about uses of the word *disagreement*, and the role of those intuitions in semantic theorizing, see Capellen and Hawthorne (2009).

[36] MacFarlane (2000).

This form of deception is far-fetched, but it is not so far-fetched as to be incomprehensible. Insofar as it is comprehensible, it is possible to imagine a listener disagreeing with the speaker's implication that she speaks the language she appears to be speaking.

Although any mode of information transfer may form the basis of a disagreement, there is variation in the options available to the listener for responding. One useful feature of the spectrum described above is that it provides a backdrop against which those options can be mapped. For example, if Bobby were to feel mislead by Alphie's assertion in (1), he can respond with (6):

(6)   Nuh uh. Only ten voted against the stimulus!

By contrast, if he feels mislead by the geographical implications of Alphie's phonological patterns, he cannot felicitously respond with (7):

(7)   #Nuh uh. You're from Canberra!

More generally, if a listener objects to information from the far right end of the spectrum, her only real option is to halt the conversation, and introduce her objection as a new topic.

(8)   Hold on. Forget politics for a minute—why are you talking like that?
        You're from Canberra!

There is a lot to say on the topic of topics and the discourse effects and requirements of introducing new ones.[37] What is clear however, is that there is no single conversational move available to Bobby that would seamlessly carry forward the conversation while allowing him to voice his objection to Alphie's presenting himself as a speaker of American English.

At the ends of the spectrum then, the difference is clear. A listener who disagrees with the propositional content of the expression uttered can simply deny what was said, availing herself of a range of distinctive and undisruptive linguistic tools: *nuh uh, no they didn't, nope,* etc. This form of response is what I will refer to simply as *linguistic denial.* That response is unavailable in the case of disagreements over information at the other

---

[37] See Roberts (1996) for an overview.

end of the spectrum. A listener cannot employ ordinary linguistic denial while nevertheless making it clear that the information she targets is phonological, or connotative, or bronchitis-related, etc. The question then is what happens along other portions of the spectrum. Are there options between normal linguistic denial and the full-scale introduction of a new topic? Where are the boundaries?

4.   Metalinguistic denials

Objections to the propositional content of an utterance can be raised by means of ordinary linguistic denial while objections to information at the opposing end of the spectrum can be raised only by introducing the objection as a new topic. However, these two alternatives—linguistic denial or a break in the conversation—are not exhaustive. Horn (1989) describes an additional option he calls *metalinguistic negation*[38], a marked reading of the negation operator on which it targets some aspect of the utterance other than the proposition expressed. For symmetry with *linguistic denial,* I will diverge in nomenclature from Horn, by referring to the phenomenon as *metalinguistic denial,* rather than *metalinguistic negation.* For reasons that will become apparent, I believe this terminology does little violence to Horn's account, and is in some ways more appropriate.

In the first epigraph to his chapter on the subject, Horn cites Jespersen (1933) giving an early description of this exceptional reading.

> With quantitative terms, *not* nearly always means 'less than'…but exceptionally these combinations [*not once, not much, not three, not half full*] may convey another meaning; this is the case if we stress the word following *not* and give it the peculiar intonation indicative of contradiction, and especially, if the negation is followed by a more exact indication; not <u>lukewarm</u>, but really hot; not <u>once</u>, but two or three times, etc.[39]

In the second epigraph, Horn concisely makes the point that the phenomenon Jespersen observes goes well beyond expressions containing quantitative terms like *much* or *three*:

---

[38] The name itself comes from Ducrot (1972).

[39] Jespersen (1933), pp. 300-301. Cited in Horn (1989), p.362.

"It's not a car, it's a Volkswagen."[40] What is this marked reading, how is it distinguished from ordinary uses of the negation, and of the many kinds of information conveyed by an utterance, what are its possible targets? I'll consider these questions in turn.

In addition to the examples cited in his epigraphs, Horn provides a large number of expressions exemplifying metalinguistic denial. To give a flavor of the phenomenon he has in mind, here are just a few of those examples.

(9)   The king of France is not bald—because there is no king of France.

(10)   *Some* men aren't chauvinists—*all* men are chauvinists.

(11)   Chris didn't *manage* to solve the problems—he solved them easily.

(12)   I didn't trap two mon*geese*—I trapped two mon*gooses*.

(13)   It's not stewed bunny, honey, it's *civet de lapin*.

(14)   *I'm* not *his* brother—*he's my* brother.

According to Horn, what makes expressions like (9)-(14) possible is a pragmatic ambiguity in speakers' usage of the negation operator. The usage in (9)-(14) serves not to negate the propositions literally expressed in the targeted assertion, but rather the acceptability of having made the assertion in the first place. The proper analysis of (10), for example, would thus be

(15)   TRUE BUT NOT ASSERTABLE(Some men are chauvinists); All men are chauvinists.

The proposition expressed by *some men are chauvinists* is true, according to the speaker of (10). The problem with an utterance of *some men are chauvinists* is that it has a false implicature. By the maxim of quantity, a listener could be expected to infer that it is not the case that all men are chauvinists. If that were the case, after all, the speaker would have said so. Since the implicature is false, it is not acceptable to utter an expression implicating it. The speaker of (10) uses the negation to target that acceptability in the first clause of her denial, and then in the second, contributes the more accurate statement.

---

[40] VW commercial and advertisement. Cited in Horn (*ibid)*.

A number of other philosophers[41] have proposed a full-blown semantic ambiguity in the negation operator in response to expressions like that in (9). Horn's analysis differs from those in that, on his view, the operator itself has a single meaning, while speakers exhibit a "duality of use" as they use and interpret the negation, not just in expressions like (9), but in the full range of expressions exemplified by (9)-(14).[42] Setting aside the question of whether paradigm cases of metalinguistic uses of the negation should be analyzed as a semantic or a pragmatic ambiguity, it is possible to step back and ask if the full range of cases constitute a unified phenomenon. In other words, do metalinguistic denials make for a natural class, and if so, how are they to be distinguished, when proof is called for, from ordinary cases of linguistic denial? Horn describes a number of tests for metalinguistic denial—behaviors that unify the various instances and that set them apart from ordinary linguistic denial. I'll focus on two of those tests here: failure to license polarity items, and inability to embed morphologically.

Ordinarily, the positive polarity items *some, already,* and *somewhat* do not occur under negations. In a typical negated expression, they are replaced with corresponding negative polarity items such as *any, yet,* and *at all.* However, Horn observes that metalinguistic negations do not trigger this behavior. Consider the following pairs:

(16) (a) John didn't see *some* of his students at the bar—he saw *all* of them.

    (b) #John didn't see *any* of his students at the bar—he saw *all* of them.

(17) (a) The king of France isn't *already* bald—there is no king of France!

    (b) #The king of France isn't *yet* bald—there is no king of France!

(18) (a) John wasn't *somewhat* worried; he was worried sick!

    (b) #John wasn't worried *at all*; he was worried sick!

In the cases Horn classifies as metalinguistic negation, the ordinary pattern is reversed. Furthermore, the pattern does not depend on strict pairing of positive and negative

_____

[41] See Horn (1989), p.369 for an overview of the various views.

[42] Horn's terminology—metalinguistic *negation*—might be taken to imply that there are two different kinds of negation. That, of course, is precisely the view he advocates against. Horn's view is that speakers can use the negation in two different ways in their actual practice of denying things. For that reason, I feel justified in making the terminological switch from metalinguistic *negation* to metalinguistic *denial.* Nothing profound hangs on this terminological choice.

polarity items. Unlike *some, already,* and *somewhat,* the expression *always* can occur easily under a negation, as in (19).

(19)   She hasn't always been happy; she used to be sad.

 Nevertheless, *always* can be replaced by the NPI *ever,* provided the logically stronger meaning is acceptable and that *ever* occurs in the range of an NPI licensing expression.

(20)   She hasn't ever been happy; she used to be sad—now she's downright depressed.

But when the denial is metalinguistic, the NPI *ever* is infelicitous even under the negation.

(21)   (in response to "She's always been happy")

    (a)   She hasn't always been happy; she's always been *ecstatic.*

    (b)   #She hasn't ever been happy; she's always been *ecstatic*.

This pattern provides a useful test for distinguishing metalinguistic from ordinary denials. If NPI's are licensed in the usual way by a negation, then the denial in which they appear is an ordinary linguistic denial. If they are unlicensed even under a negation, then the denial in which they appear is metalinguistic.

Another of Horn's tests concerns the ability or inability of negation to appear morphologically embedded with another expression in the sentence. In descriptive denials, negation can appear either embedded or as a distinct lexical item.[43]

(22)   (a)   It's possible that John is in London.

    (b)   Nuh uh, it's not possible that John is in London.

    (c)   Nuh uh, it's impossible that John is in London.

(23)   (a)   Bobby is happy.

    (b)   Nuh uh, Bobby is not happy.

    (c)   Nuh uh, Bobby is unhappy.

(24)   (a)   Ross is a contender for heavy-weight champion.

    (b)   Nuh uh, Ross isn't a contender.

---

[43] This is not to say that the meaning of the negation is identical between embedded and non-embedded occurrences. A person might fail to qualify as *happy* without thereby qualifying as *unhappy*. The point remains however, that either form is available for use in a descriptive denial, even if the denials themselves make subtly different contributions to the conversational context.

(c)   Nuh uh, Ross is a non-contender.

In the case of metalinguistic negations, however, a morphologically incorporated
negation renders the desired reading unavailable.

        (25)   (a)   It's possible that John is in London.

                (b)   It's not *possible* that John is in London; it's downright necessary!

                (c)   #It's *impossible* that John is in London; it's downright necessary!

        (26)   (a)   Bobby is happy.

                (b)   Bobby isn't *happy*; he's ecstatic!

                (c)   #Bobby is *unhappy*; he's ecstatic!

        (27)   (a)   Ross is a contender.

                (b)   Ross is not a *contender*; he's a shoo-in!

                (c)   #Ross is a *non-contender*; he's a shoo-in!

Though the distinct lexical item *not* and the embedded forms *iN-, un-,* and *non-* typically
make similar contributions to the meaning of the expression, in the case of metalinguistic
denials, they come widely apart[44]. When the negation appears as a distinct lexical item,
either reading is available: the expression "It's not possible that John is in London" could
be interpreted either as in (22b) or as in (25b), depending on intonation. By contrast,
morphologically embedded negations can only be interpreted descriptively.

        Horn argues that this striking feature of metalinguistic denials, along with the
failure to trigger normal polarity behavior, are both explained by the negation's operating
"on another level" from the rest of the expression.

> Whenever negation is used metalinguistically to deny the appropriateness of using
> a predicate which would yield a true but misleading assertion (one which would
> induce false conventional or conversational implicata), it operates on another
> level from the rest of the clause in which it is superficially situated, whence its
> impotence to trigger polarity items with that clause […]. For the same reason, the
> metalinguistic operator cannot incorporate morphologically as the *un-* or *iN-*
> prefix…. [T]he incorporation of a negative element as a prefix on the following

_____

[44] There is a small range of exceptions to this generalization which will not matter for my purposes. For a
discussion see Horn (*ibid*), p. 395-397.

element is only possible when those two elements are immediate constituents within a single level of analysis.[45]

If a negation can occur as a morphologically embedded component of another expression, then the denial in which they appear is a linguistic denial. If the negation cannot appear as a morphologically embedded component of another expression without forcing an undesired reading or rendering the whole expression infelicitous, then the denial is a metalinguistic denial.

These two generalizations about metalinguistic denials—the failure of the negations to license polarity items, and the inability of negations to morphologically embed—unify the the denials Horn characterizes as metalinguistic, and they provide diagnostics that can be applied in cases where the nature of a particular denial is unclear. If polarity items behave in the usual way, and if the negation can embed without rendering the intended meaning unavailable, then the negation in question is descriptive. If the behavior of polarity items is reversed relative to their typical behavior, and if the negation cannot embed, then the negation in question is metalinguistic.

Horn himself is concerned with denials targeting familiar cases of conventional or conversational implicatures, as well as examples of false presupposition and objectionable manners of speaking—mispronunciations and misleading or objectionable connotations. Each of these categories is, in the terminology introduced above, a mode of information transfer, situated somewhere along the spectrum described in Section 1. What Horn has demonstrated is that, between the far left end of that spectrum, where linguistic denials are available, and the far right end, where only the introduction of a new conversational topic can serve as an objection, there is a rather vast middle ground. In that middle ground, where information is communicated via a mix of linguistic and nonlinguistic evidence—via presupposition, implicatures, connotations, and the like— speakers can make use of this alternative, marked reading of the negation to object and to draw attention to the fact that they object not to propositional content but to some other feature of the utterance.

---

[45] *ibid.* pp. 392-393.

What then about the very center of the spectrum? Horn does not consider disagreements about information of the kind described by Ludlow and Barker. To return to the case of Alphie and Bobby's discussion of Secretariat, suppose that Bobby deduces, accurately, that John uses *athlete* to refer both to humans and to non-human animals that engage in sport. Let's go on to suppose that Alphie finds such a usage objectionable. Or suppose that John is asked what counts as tall around here and responds with his utterance about Feynman, but that Bobby, another participant in the conversation, believes that in fact the salient standard of tallness does *not* qualify Feynman as tall. Let us call disagreements over character-sharpening *character disagreements,* and disagreements over context-sharpening *context disagreements.* Just as the transfer of information along those two channels can be considered *metalinguistic* usages, let us refer to disagreements over information transferred in this way as *metalinguistic disagreements.* The term seems to fit. What could be more metalinguistic than a disagreement over the appropriate usage of a term, or a disagreement over how to set the contextually sensitive parameters of context-sensitive terminology? Strikingly, Horn's tests demonstrate that metalinguistic denials are unavailable in metalinguistic disagreements.

5.   Metalinguistic Disagreements Do Not license Metalinguistic Denial

Consider first Horn's morphological incorporation test. According to that test, if a negation can appear embedded with another part of the expression, then the denial in which they appear is descriptive. If embedding renders the expression infelicitous, or forces an undesired reading, then the denial is metalinguistic. So consider the *athlete* debate, and suppose that Bobby does after all decide to put up a fight, just like the participants in the original talk radio show observed by Ludlow.

(28)   (a)   Secretariat is an athlete.

         (b)   Secretariat isn't an athlete!

         (c)   Secretariat is a non-athlete!

(28c) is less natural than (28b), but it's not much worse than (24c) (repeated as (29c)) and it's substantially better than (27c) (repeated here as (30c)). As a set, the pattern in (28) is much closer to (29) than to (30).

(29) (a) Ross is a contender for heavy-weight champion.

(b) Nuh uh, Ross isn't a contended.

(c) Nuh uh, Ross is a non-contender.

(30) (a) Ross is a contender.

(b) Ross is not a *contender*; he's a shoo-in!

(c) #Ross is a *non-contender*; he's a shoo-in!

The point becomes even clearer if we move to other examples. Consider the widespread debate about Pluto's status as a planet. That issue has sparked a huge amount of interest in popular culture, despite the fact that even ordinary speakers seem well aware that what they are debating is no deep metaphysical fact but rather a matter of usage and the social cache that follows from certain labels.[46] So imagine a conversation between John, the man on the street and Pluto lover, and Sally, the astronomer.

(31) John: Pluto is a planet.

(32) Sally: Nuh uh, Pluto is not a planet; it's just a large asteroid.

In this instance, it's plausible to imagine that both John and Sally are well aware that what they are arguing about is a matter of usage. Does *planet* mean what John means by it—a collection of nine special objects in the sky, memorizable with the mnemonic "My Very Educated Mother Just Serves Us Nasty Pickles", one of which happens to be endearingly small and quirky? Or does it mean what Sally and her colleagues deem it to mean in light of their studies of the actual mass, size, and orbits of Pluto and similar astronomical bodies?[47] This seems like prime territory for metalinguistic negation. After all, Sally might well accept that on the ordinary English usage of the term, Pluto is a

---

[46] See Tyson (2009) for an overview of both the scientific and cultural debates.

[47] I will stipulate for now that John and Sally mean different things by *planet*. That stipulation runs counter to certain externalist theories of reference according to which John and Sally mean the same thing by *planet* no matter how different their usage. A parallel, though technically independent, concern can be raised by the similar case of difference in contextual parameter setting, considered below. I address such concerns in Section 6, below.

planet. If she does, then she'll accept that what John says is literally true. What she disputes is that that usage is the best available. It ought, then, be possible for her to issue a denial of (31) that could be analyzed as in (33).

>(33)   TRUE BUT NOT ASSERTABLE(Pluto is a planet$_{\text{ordinary English}}$); Pluto is not a planet$_{\text{astronomers}}$.

Unfortunately, the denial that would be analyzed as in (33) would be uttered as in (34).

>(34)   Pluto is not a *planet*; pluto is not a planet.

(34) itself is pretty awkward, but that awkwardness can be reduced somewhat by replacing the second clause with an alternative that entails the same thing but provides a bit more information and makes the whole expression less repetitive.

>(35)   Pluto is not a *planet;* it's just a big asteroid.

Now (35) is a reasonable enough way to respond to (31). In fact it's just (32) with a bit of emphasis on the word *planet*. The question is whether there is any reason to analyze (35) as a metalinguistic negation, as in (36),

>(36)   TRUE BUT NOT ASSERTABLE(Pluto is a planet$_{\text{ordinary English}}$); Pluto is just a big asteroid.

or simply as a descriptive denial, as in (37):

>(37)   NOT(Pluto is a planet$_{\text{astronomers}}$); Pluto is just a big asteroid.

Intuitively, I think (37) seems more promising than (36) as an analysis of (35), but, fortunately, we will not need to rely on my intuitions. If (35) were a metalinguistic denial, then the negation in it could not be morphologically embedded. But it can be morphologically embedded.

>(38)   Pluto is a *non-planet*; it's just a big asteroid.

According to Horn's incorporation test then, Sally's denial of John's utterance is not metalinguistic.

Turning to Horn's other test—the failure of metalinguistic denials to license NPI's—we find more proof that Sally's options in responding to John do not include metalinguistic denial. Suppose that John, stubbornly insisting on the primacy of his preferred usage of *planet*, utters (39).

>(39)   Pluto has always been a planet.

Sally, in responding, could very easily reply with (40).

(40)   Pluto has not always been a planet; it has always been a big asteroid![48]

If (40) were a metalinguistic denial, then the negation in the first clause would not license NPI's in the normal way. But the negation in the first clause does license NPI's in the normal way.

(41)   Pluto hasn't ever been a planet; it has always been just a big asteroid!

Horn's tests thus demonstrate that even if she wanted to, Sally could not felicitously deny John's utterance in such a way that she acknowledged its literal truth, while objecting to its assertibility on other grounds.

The same result we've seen above with character disagreement holds for context-disagreements as well. Recall Barker's case concerning the standards for tallness. *Tall,* unfortunately, does not happily accept morphologically embedded negations. (*Non-tall* is comprehensible enough, but hardly a mainstay of ordinary conversation.) Let's consider then an alternative, *adequate*. Betty is a visiting student at Bobby's philosophy department. In an attempt to get her bearings and understand what will be expected of her, she asks "What counts as adequate academic performance around here?" Bobby points to a mutual friend, Alphie, whose academic performance is well known to both him and Betty.

(42)   Well, Alphie's academic performance is adequate.

So far, the example is parallel to Barker's example. But now suppose that another party to the conversation, Jim, disagrees with the context-sharpening information that Bobby has conveyed to Betty. Jim happens to think that Alphie very much fails to satisfy the local standards of adequacy. How might he reply?

(43)   Alphie's performance is not adequate; he just barely gets by!

The issues here are subtle, and it seems excessive to expect ordinary speakers to be consciously aware of the exact nature of their conversation. Nevertheless, we might easily imagine that Jim, Bobby, and Betty all have detailed knowledge of Alphie's

---

[48] (40) happens to contain an internally cancelled implicature. Sally's utterance of *Pluto has not always been a planet* would typically be interpreted by a speaker as implicating that Pluto is now a planet. Sally cancels that implicature with her second clause. This feature of her expression is orthogonal to its relevance here. The important datum for my argument is in (41).

academic record and that each knows that other two have that knowledge as well. While it may be tacit, it is therefore reasonable to assume that the speakers understand that what is at issue here is not a question about Alphie's performance, but rather a question about the local standards of adequacy, and whether barely getting by is enough to qualify.[49] Bobby thinks it is. Jim thinks that to deserve the label *adequate*, one must perform at least a bit beyond the bare minimum. Since Jim realizes that both he and Bobby are aware of the academic facts, he might well realize that Bobby applies a different standard of adequacy. If he does, he will realize that what Bobby uttered is literally true, though defective insofar as it applies a too-lax definition of adequacy.

If it's sensible to think of Jim realizing (tacitly) that what's at issue here are the standards of adequacy—and because the facts of Alphie's performance are mutual knowledge, it is—we might then predict that Jim could deny (42) in a way that is best analyzed as in (44):

(44)   TRUE BUT NOT ASSERTABLE(Alphie's performance is adequate$_{Bobby}$); Alphie's performance is not adequate$_{Jim.}$

Unfortunately, as in the case of character disagreements, a denial that is analyzed as in (44) would be rather unnatural when actually uttered:

(45)   Alphie's performance is not *adequate*; his performance is not adequate.

Although (45) closely mirrors the paradigm cases of metalinguistic denial, the result in this case is very unnatural sounding. As in the case of (34), however, this awkward denial can be improved upon. Replacing the second clause with something more informative and less repetitive yields a denial like that in (46):

(46)   Alphie's performance is not *adequate*; he just barely gets by.

(46) is a reasonable enough way to deny to Bobby's claim, but it again raises the question of whether it is best analyzed as metalinguistic. If it is, then the negation should not be able to appear morphologically incorporated with another element, and it should not

---

license NPI's. Both of those tests demonstrate that (46) cannot be analyzed as a metalinguistic denial.

Consider first the incorporation test. With very little, if any, change in meaning, the denial in (46) can be expressed as in (47):

(47)   Alphie's performance is inadequate; he just barely gets by.

If (46) could be analyzed as a metalinguistic negation, that change should render the expression infelicitous, or force a change in meaning. But it does not. Now consider the polarity test. To make use of that test, we'll first have to lay the groundwork for the inclusion of an NPI, so suppose that, in response to Betty's question about the local standards of adequacy, Bobby replies with (48).

(48)   Well, Alphie's performance has always been adequate.

In response to (48), Jim might well reply with (49).

(49)   Alphie's performance has not always been adequate; he's always just gotten by!

But, assuming he's willing to make the logically stronger claim, (50) is an equally felicitous response.

(50)   Alphie's performance has not *ever* been adequate; he's always just gotten by!

If Jim could reply to Bobby in a manner indicating that Bobby's claim was true but defective—in other words, if metalinguistic denial were an option for Jim in his dispute with Bobby—then the NPI *ever* would not be licensed in his denial. But *ever* is licensed. Therefore, even if he wanted to, Jim could not signal his disagreement with Bobby's context-sharpening implication that *adequacy* is so defined as to include Alphie. Not, that is, without simply denying Bobby's claim flat out. Metalinguistic disagreements do not license metalinguistic denials.

6.   Explanation Attempt I: The Cases Have Been Misdescribed in Virtue of Hewing Too Closely to Speakers' Usage

One way to respond to the failure of metalinguistic disagreements to license metalinguistic negations is to take issue with the analysis according to which such

discourses *are* metalinguistic disagreements. In both the *planet* case, and in the *adequate* case, I have assumed that the literal content of the expressions involved was a function, largely, of the intentions of individual speakers. That allowed me to assume that by *planet*, John meant what he thought he meant: something corresponding closely to the use of *planet* in ordinary English. Likewise, it allowed me to assume that in his usage of *adequate*, Bobby fixed the contextually variable parameters in the way that he thought he did, a setting according to which Alphie qualified as adequate.

Those assumptions are not entirely uncontroversial. Many philosophers[50] take the view that there can be a substantial disconnect between a speaker's usage and intentions on the one hand, and the content they actually express on the other. If that disconnect is substantial enough, then we could imagine an analysis of the above cases according to which, whatever the speakers *thought* they were saying, they *in fact* expressed propositions that directly contradicted the propositions expressed by their interlocutors. Whatever John intended with his use of *planet*, the content of his expression was fixed by, say, the concept of *planet* described in astronomy. However Alphie thought he was fixing the threshold of adequacy, the contextual parameters of his words were in fact fixed by, say, Jim's views on the matter.[51]

These issues clearly interact with entrenched and ongoing debates between *internalists* and *externalists* about meaning. I'll put off for now the task of parsing out the precise nature of that interaction; nevertheless I will refer to the proposed alternative as an *externalist* analysis, without thereby intending to attribute it to any particular externalist. The alternative analysis is externalist in the very simple sense that what John and Betty meant by their word *planet*, for example, is determined largely by factors

---

[50] See Putnam (1975), Burge (1979), or Lewis (1983), among others. Views about the role of speaker intentions with respect to the extension of a term need not line up with views about the role of speaker intentions with respect to fixing contextual parameters. The arguments I will give in this section apply to both, but some philosophers who take an externalist line about reference fixing will be happy to grant the point when it comes to fixing context. See, for example, Stalnaker (1970). Thanks to Andy Egan for helping to clarify this point.

[51] Whatever factors outside of John and Bobby's acquaintance or control are doing the meaning determination here, it need not be the case that the actual meanings conform to what Betty and Alphie think they meant by *their* words. The meaning determining factors may settle everyone's choice of concept or parameter settings on some third alternative.

outside of their acquaintance or control; the correct semantic theory of their words must be sensitive to those external factors in such a way that the two speakers are correctly described as meaning the same thing no matter the systematicity of variance in their usage and intentions. A sufficiently strong and wide-ranging form of externalism would conclude that the conflicts between John and Betty and between Alphie and Jim represent ordinary disagreements over content. If that is the case, then there is no mystery to the fact that they fail to license metalinguistic denials.

For a variety of reasons, this line of response cannot work. First, the proposed alternative analysis of these cases is not plausible. The problem is that in any externalist theory, a certain disconnect is posited between speaker usage and semantic theory, and in these particular cases that disconnect is too large to be believed.[52] Certainly, no theory of literal content can hew perfectly to the usage of ordinary speakers. But to posit a massive disconnect between how a speaker talks and the proper theory of their words is nevertheless a cost. That cost is worth paying when it comes to factoring out pragmatic factors in language use, for example, or when it comes to easily-controlled-for speaker error. We know how to adjust a semantic theory of *some* to allow for the fact that speakers often use it to mean *some but not all.* We know how to adjust a semantic theory of *cat* to allow for the fact that speakers in a darkened room may use that term to describe a possum. Those costs are worth bearing and are easily paid in the currency of, for example, an independently motivated theory of pragmatics.

By contrast, the cost of the conclusion that John uses *planet,* despite his consistent and sincere assertions to the contrary, to pick out a concept excluding Pluto, is steep. If the meaning of ordinary words like *planet* is chained to homophonous scientific terminology in this way, then vast swaths of ordinary speech, speech that occurs under a range of circumstances large enough to cull out even the most subtle of intervening pragmatic or observational factors, must be categorized as false. In many of those circumstances, speakers like John will have no particular interest in or need for deference to the scientific experts, nor will the metaphysical naturalness sought by those experts be

---

[52] See Chapter 3 of this thesis for a more developed version of this argument.

a constraint of any special interest on the appropriateness of John's concepts. And of course, the case of Pluto is not unusual, except in the degree to which the usage of experts has caught the interest of the popular press. A semantic theory that, in deference to physicists, branded false the optimist's sentiment that "love is a force for good" (how could it be if it's not a force?); or that, in deference to Chomsky's notion of i-language, labelled false the ordinary observation that "English is a language" would be no more plausible.

The other problem with this line of response is that whatever the status of these individual examples, only the most implausible externalism could deny that there are *some* cases that are correctly described as metalinguistic disagreement. However strong a force external factors provide in assuring that meanings are stable across speakers, there must remain *some* cases of disputes about usage, or proposed reference shift, or polysenym preference, or contextual parameter selection of the kind that generate metalinguistic disagreements. Even those who give pride of place to notions like "linguistic community" or "social practice" over the usage of individual speakers will face situations where speakers from different but overlapping communities speak to one another in mutually comprehensible but non-identical dialects. When such speakers find themselves objecting to one another's choice of words—an undeniable feature of real-life language use—they will disagree in precisely the kind of non-metalinguistic-denial-licensing ways described above.

(51)　Southerner:

That Sprite is a Coke.

(52)　Northerner:

(a)　That Sprite isn't a Coke; it's a Sprite!

(b)　That Sprite is a non-Coke; it's a Sprite!

(c)　That Sprite hasn't ever been a Coke; it's always been a Sprite!

(d)　#That Sprite isn't a *Coke*; it's not a *Coke*. [TRUE BUT NOT ASSERTABLE(That Sprite is a Coke$_{Southern}$); That Sprite is not a Coke$_{Northern}$.]

Only a form of externalism that denied the very possibility of language change or variation could entirely avoid the problem of how to deny information communicated metalinguistically. There are indeed cases of metalinguistic disagreement, and they do fail to license metalinguistic denial, and that fact remains in need of explanation.

7.   Explanation Attempt II, Part I: R-Based Implicatures Don't License
        Metalinguistic Negation Either

        Fortunately, such an alternative explanation is forthcoming. The alternative explanation works by demonstrating that there is a logical parallelism between metalinguistic disagreements and another kind of non-content disagreement that also fails to license metalinguistic denial, namely relevance, or *R-based* implicatures.

        In his discussion of implicatures and negation, Horn observes a difference between implicatures that depend on the Maxim of Quantity (*Q-based* implicatures), and implicatures that depend on the Maxim of Relevance, or *R-based* implicatures. Horn points out that a speaker taking issue with information transmitted via quantity implicature can respond with a metalinguistic denial, but that a speaker objecting to information transmitted via relevance implicature cannot respond with a metalinguistic denial. Consider first a straightforward case of quantity implicature. Suppose Alphie utters (53).

        (53)   Sally is smart.

In a typical context, (53) could reasonably be taken to implicate that Sally is not a genius. The reason is that on the scale of intelligence, *genius* is stronger than *smart.* If Alphie had known that Sally was a genius, it would have been uncooperative of him to utter (53) in virtue of the fact that he could have said something logically stronger. Alphie's interlocutor Betty assumes that John is being cooperative, and therefore takes him to have communicated the negation of that stronger claim.

        Now if Betty herself happens to believe that the stronger claim is true, she may reply as in (54).

        (54)   Sally isn't *smart*; she's a genius!

Betty doesn't deny that Sally is smart. The negation in the first clause of her expression targets the assertibility of Alphie's utterance in virtue of its false implicature. It is a straightforward case of metalinguistic denial used to target information beyond the literally expressed content of an utterance.

Now consider a case of relevance implicature. Suppose Alphie utters (55).

(55)   Sally was able to solve the problem.

In typical contexts, an utterance of (55) can be taken to implicate that Sally indeed *solved* the problem, even if the proposition literally expressed is the weaker claim that she had the requisite ability. The reason the stronger claim is implicated is that in many of those contexts, what matters for purposes of the conversation is whether the problem was solved and who solved it. If John had meant to communicate merely the fact that Sally had the ability to solve the problem, without contributing to the conversational goal of establishing whether the problem was solved and by whom, then his utterance would have been uncooperative in virtue of expressing an irrelevant proposition. Betty assumes that John is being cooperative, and therefore takes him to have communicated the stronger claim.

Now if Betty agrees that Sally had the ability to solve the problem, but believes that the implicated proposition—that Sally actually solved it—is false, she may reply as in (56), but not as in (57), where (57) is interpreted as in (58).

(56)   Nuh uh, Sally didn't solve the problem!

(57)   #Sally wasn't *able* to solve the problem; she didn't solve it!

(58)   TRUE BUT NOT ASSERTABLE(Sally was able to solve the problem);
         Sally didn't solve the problem.

On a close reading, we can see that Betty's response in (56) does not actually refute what Alphie expressed in (55). Nevertheless, the response in (56) is vastly more natural than the entirely infelicitous attempt at metalinguistic denial in (57). This despite the fact that (57), as interpreted in (58), is structurally analogous to the metalinguistic denial in (54). Even if Betty's stance towards (55) is just the same as it was towards (53)—agreement with the literal content and disagreement with the implicature—her options are different.

Disagreements over quantity implicatures license metalinguistic denials but disagreements over relevance implicatures do not.

What to make of the difference? Horn argues that the difference in licensing behaviors is due to a feature of the logical relationships among the propositions literally expressed and implicated. Consider again the utterance generating a quantity implicature, repeated here as (59), along with the stronger proposition in (60).

(59)   Sally is smart.

(60)   Sally is a genius.

Horn observes that the implicature of an utterance of (59) is *the negation* of the (60). The proposition in (60) itself unilaterally entails the proposition in (59). With those facts in mind, let us reconsider Betty's utterance of (54), repeated here as (61).

(61)   Sally isn't *smart*; she's a genius!

(61), like all standard metalinguistic denials[53], contains two components. I'll refer to them as the *denial component*—"Sally isn't smart"—and the *correction component* —"Sally is a genius". Horn observes that if (61) were interpreted descriptively, such that the negation operator in the denial component targeted the proposition that Sally is smart, then the denial and the correction components of Betty's expression would be mutually inconsistent. If Sally isn't smart then she can't be a genius.

Horn argues that this is a crucial feature of metalinguistic negations. What he observes is that metalinguistic negations are only possible in situations where, on a descriptive reading, the denial and correction components are mutually inconsistent. That this is true across the range of metalinguistic denials can be seen by recalling some of the examples with which we began:

(62)   *Some* men aren't chauvinists—*all* men are chauvinists!

(63)   It's not *stewed bunny*; it's *civet de lapin*.

(64)   *I'm* not *his* brother—*he's my* brother.

In each of these cases, the denial component of the sequence would be logically inconsistent with the correction component, were the negation in the denial component

---

[53] Portions of a metalinguistic negation can sometimes be elided. That fact won't be relevant here.

interpreted in the ordinary way. If it is not the case that some men are chauvinists, then it can't be the case that all men are chauvinists; if the dish isn't stewed bunny, then it can't be *civet de lapin*; and if I am not his brother, then it had better not be the case that he is my brother.

Horn's explanation of this pattern turns on a hypothesis about language parsing. His idea is that listeners assume as a default that negations are interpreted descriptively, but that when the inconsistency is discovered, alternative interpretations are considered and rejected until a coherent interpretation is discovered.

> It might even be speculated (although I shall not try to support this speculation here) that there is an inherent ordering within the processing of metalinguistic negation: the addressee, on recognizing that a given negation cannot be coherently read descriptively (as denying a given predicate of a given subject or as an internal, predicate term negation), will try first to take it as a rejection of the conventional implicata (or presuppositions) associated with the negated utterance, then (if that fails) as a rejection of its potential conversational implicata, then (if that fails) as a rejection of the formal properties (grammar or phonology) of that utterance.)[54]

On Horn's story, the inconsistency between the denial and correction components of the sequence forces the (charitable) listener to go back and reinterpret the denial component of the sequence in such a way that the speaker can be interpreted as having expressed something non-contradictory. By reinterpreting the negation as targeting some feature of the original expression other than its literal content, that more charitable reading is made possible.

Whether or not Horn's reparsing explanation is the correct analysis of the pattern —an issue that, with Horn, I'm happy to leave to the psycholinguists—the generalization itself holds perfectly in the cases we've considered. Recall our case of an utterance generating a relevance implicature, repeated in (65), along with the implicated proposition in (66).

(65)   Sally was able to solve the problem.

(66)   Sally solved the problem.

---

[54] Horn (*ibid.*), p. 391.

In the case of quantity implicature, the expression of a weak proposition implicated *the negation* of a strong proposition: *Sally is smart* implicated that Sally is *not* a genius. A listener objecting to that implicature therefore *believes* the strong proposition—that Sally is a genius—and is therefore committed to the truth of the weaker proposition. In the case of relevance implicature by contrast, the expression of the weak proposition implicates the strong proposition itself, not its negation. *Sally was able to solve the problem* implicates that she solved it. So a speaker who *disputes* a relevance implicature disputes something stronger than the proposition literally expressed. But disputing something stronger than the proposition literally expressed does not generate any commitment with respect to that weaker proposition: that Sally didn't solve the problem is independent of her ability to solve it.

Returning to the kind of metalinguistic denial this dispute would generate, we can see that the sequence has precisely the form that, according to Horn, precludes a metalinguistic reading:

(67)   #Sally wasn't able to solve the problem; she didn't solve it!

According to Horn's generalization, a metalinguistic denial must contain denial and correction components that, on a descriptive reading of the denial component, are logically inconsistent: If Sally isn't smart then she can't be a genius, so if somebody says she isn't smart but she is a genius, they must have intended a marked reading. In (67), by contrast, the two components are entirely consistent even on a descriptive reading of the denial component. Sally's being unable to solve the problem is consistent with her failure to solve it, and the listener parsing an utterance of (67) will therefore have no reason to go back and reinterpret the negation in its first clause. So even if Betty agrees with the literal content of Alphie's utterance and disagrees only with the implicature, she cannot use metalinguistic denial to express that set of views.

What Horn's observation demonstrates is that even in clear cut examples of implicature, there will be some disagreements in which metalinguistic denials are unlicensed. In particular, R-based implicatures, which are very much cases of implicature, will not license metalinguistic denials. They fail to do so because the denial

component of such a sequence would be consistent with the correction component. Since Horn's observation makes no reference to other unique features of R-based implicatures, we can assume it to generalize beyond relevance implicatures. The observation relies on a certain logical relationship between the expression that would be used to deny a previous utterance, and the expression that would be used to correct it. Any case of non-content disagreement in which that logical relationship obtains, and in which metalinguistic denials are unlicensed, thus patterns *with* implicatures, despite any initial oddness associated with the lack of metalinguistic denial. As we'll see, that's precisely the state of things with disputes over sharpening.

8.   Explanation Attempt II, Part II: The Logical Relationships that Horn Observes in the Case of R-based Implicatures Also Hold in the Case of Sharpening Uses

Horn's generalization states that metalinguistic denials are possible only in cases where, on a descriptive reading of the denial component, the denial and correction components of the sequence are inconsistent.[55] Whether his reparsing-explanation of that generalization is accurate or not, his generalization accurately characterizes what otherwise is a rather odd disconnect among various forms of implicature. The generalization becomes even more useful once we realize that, properly amended, it carries over perfectly to explain the absence of metalinguistic denials in cases of metalinguistic disagreements.

To see how Horn's generalization carries over to metalinguistic disagreements, recall the dispute between John, the man on the street and Betty, the astronomer. John utters the expression in (68), expressing his belief that Pluto is a *planet* in the ordinary sense of that word.

(68)   Pluto is a planet.

We imagined, plausibly, that Betty the astronomer agrees that Pluto falls in the extension of *planet* as it is used by non-astronomers, but that she takes issue with that usage,

_____

[55] Geurts (1998) raises potential counterexamples to this generalization. I think that there is a response to be made on Horn's behalf, but I will not be able to address the issue here.

astronomers having discovered that Pluto has less in common Venus, Saturn, etc. and more in common with other large asteroids. This set of views is analogous to the cases above, in which Betty agrees with the literal content of Alphie's utterances—that Sally is smart; that Sally had the ability to solve the problem—but disagrees with the implicatures —that Sally is not a genius; that Sally solved the problem.

To see what kind of responses are available to Betty, we imagined that she would want to utter a denial that could be analyzed as in (69):

(69)   TRUE BUT NOT ASSERTABLE(Pluto is not a planet_{ordinary English}); it's just a large asteroid.

Sure enough, Betty can respond to John with an utterance of (70):

(70)   Pluto is not a *planet*; it's just a large asteroid.

However, on inspection, the denial component of (70) proves not to be metalinguistic but rather an ordinary descriptive denial.

So what of Horn's generalization? Well consider the two components of Betty's denial. *If* her denial were metalinguistic, then the denial component must contain the word *planet* as John uses it, since it is that expression that Betty considers unassertable. The sequence would thus look like (71):

(71)   Pluto is not a *planet*_{ordinary English}; Pluto is just a large asteroid.

When the *not* in the first clause of Betty's denial is interpreted descriptively, her claim is simply that Pluto does not fall in the extension of the ordinary English term *planet*. That proposition, whatever its actual truth value, is clearly consistent with the proposition that Pluto is just a large asteroid.

Perhaps we've moved too quickly however. Horn's explanation for the patterning of metalinguistic denial depends on the listener's interpretation of the speaker's utterance. But unlike the ordinary implicatures that Horn considers, cases of metalinguistic disagreement turn on expressions whose interpretation is at issue. So not only is the interpretation of the negation in the denial at issue, but also the interpretation of the original, objectionable expression. In the Pluto case specifically, we've assumed that what's expressed in the denial component is that Pluto is not a planet in the ordinary

English sense. But perhaps John, like Betty, knows that they use their words differently.
When he hears Betty use the word *planet,* he may then interpret her expression as in (72).

(72)   Pluto is not a *planet*astronomers; it's just a large asteroid.

Even if this were John's reading of Betty's utterance, however, it is clear that the two
components of the sequence are mutually consistent. There is thus no interpretation of
Betty's utterance on which the two components of the sequence are inconsistent. Nothing
then requires John, however he interprets the word *planet*, to reinterpret the negation in
Betty's utterance metalinguistically. Metalinguistic disagreements are thus similar to
disagreements over R-based implicatures: in both cases, attempts at metalinguistic denial
result in denial/correction sequences that are mutually consistent even on a descriptive
reading of the denial. Metalinguistic denials are thus impossible for both.

9.   Information Carried Metalinguistically Is Like Implicature in Other Ways

We've seen that from a single utterance, a listener can infer a wide range of
information. That information includes the usual suspects, like the proposition expressed,
along with entailments, presuppositions, and implicatures. But it also includes
information about the language used in the utterance, information transferred via the very
usage of that language. A listener can learn from an utterance not only about an object's
height, but also about the prevailing standards for *tallness*. She can learn not only about
athletics, but also about the preferred usage of the predicate *athlete*. She can learn not
only about astronomy, but also about the changing scientific definition of *planet.*

We've also seen that across the full range of information conveyed, interlocutors
can be every bit as disagreeable as they are with respect to the proposition expressed.
Just as nothing prevents a speaker from conveying information via channels outside of
the proposition expressed, nothing prevents a listener from taking issue with information
thereby conveyed. As Horn demonstrates, in many cases of disagreements over
information other than the proposition expressed, listeners have available to them a
device—metalinguistic denial—that makes clear the fact that the objectionable

information is not the literal content. With that in mind, it is at first quite odd that in a particular range of disagreements, the disagreements I've called *metalinguistic*—that device is unavailable.

On first pass then, and given the prevalence of externalist views about meaning among philosophers of language, one might be tempted to take that oddness as a sign that in fact those disagreements are not metalinguistic, but rather descriptive, thanks to some factor external to the speakers or their intentions that acts to keep meaning stable between them. However, that view is not only problematic on its merits, given the extreme form of externalism that would be required to avoid the problem, but it is also unnecessary. As Horn himself observes, there are other, more familiar forms of dispute that also concern aspects of meaning other than propositional content and that also fail to license metalinguistic denials—namely disagreements over relevance implicatures. And as we observed subsequently, the feature of disagreements over relevance implicatures that prevents them from licensing metalinguistic denial are shared by metalinguistic disagreements. There is therefore every reason to believe that disagreements at the center of the spectrum exist, that they may indeed be common, but that they are all too likely to fly under the radar of philosophers of language, given that the tool best suited to sussing out such non-propositional-content-focused disagreements is unavailable.

Having argued for this general understanding of the landscape, I'd like to conclude by suggesting that the similarity between metalinguistic disagreements and disagreements over implicatures is symptomatic of a more general parallelism between implicature and metalinguistic usage. Recall our original cases of metalinguistic usage: Barker-style cases of metalinguistically conveyed information about contextual parameters, and Ludlow-style cases of metalinguistically conveyed information about the extensions of predicates. In each case, we imagined a listener observing the utterance and then deducing, on the basis of their knowledge of both linguistic meanings and the facts around them, certain conclusions about the way words are to be used. But how exactly do the listeners arrive at those conclusions?

Consider one last time our original example of metalinguistic usage, the discussion between Alphie and Bobby about secretariat. To recall, Alphie utters the expression in (2), repeated here as (73).

(73)   Secretariat was an athlete.

The relevant physical characteristics of Secretariat are mutual knowledge between Alphie and Bobby, and thus Bobby, who uses *athlete* only to refer to sporting humans, is confronted with a choice. He can either conclude that Alphie has said something radically defective, or he can conclude that Alphie uses *athlete* differently. Finding the second option much more plausible, Bobby must decide whether to adopt that usage. In some cases, he will put up a fight, and will issue the kind of denials we considered in detail above. But in many cases he will not put up a fight, and will instead choose to adopt the assumption that *athlete* is best used under the present circumstances in the way that Alphie uses it.[56]

The first step in Bobby's reasoning is that Alphie uses the word *athlete* differently than he does. But that is not end of this line of reasoning. As we've seen, Bobby goes on to conclude, in some though not all cases, that *athlete* is *best* used in the way that Alphie uses it.[57] In other words, Bobby starts with the intermediate conclusion that Bobby uses *athlete* in a certain way, and reasons from there to the conclusion that *athlete* is *most appropriately* used in that way. It is that further conclusion that, on occasion, prompts Bobby to modify his own usage, and it is an objection to this conclusion that might lead Bobby to take issue with Alphie's usage. It is the analogous conclusion that allows the listener in Barker-style cases to conclude that, not only does the speaker take *tallness* to be relative to a given threshold, but that she herself ought to take it the same way. And it is an objection to that analogous conclusion that may lead the listener in a Barker-style case to take issue with the speaker's usage.

---

[56] In analogous Barker-style cases, he will adopt the assumption that the contextual parameters of tall are best set in the way that Alphie sets them.

[57] In some, perhaps many, situations, Alphie's determination that Bobby's usage is most appropriate will be for the sake of the conversation. He may indeed believe that there is some better usage available, but that it is not worth putting up the fight. Such acceptance for the sake of conversation is entirely compatible with the view I describe here.

How is this second conclusion justified? How does Bobby move from the conclusion that Alphie speaks in a certain way to the conclusion that one *ought* to speak in that way? I think the most natural answer to the question supposes Bobby to make certain assumptions about his conversational partner, assumptions according to which the partner uses his words in a manner that is most appropriate to the conversation. Bobby's assumption is thus something like what, for lack of a better name, I will call the *Metalinguistic Maxim*:

> MM: The speaker is employing the precisifications/contextual parameter settings/restrictions/polysenyms/homophones/etc. that are most appropriate to the conversation.

If Bobby's assumption is accurate, then from his conclusions about the way Alphie is actually speaking, he is licensed to draw inferences about the ways in which it is appropriate to speak. It is thus this assumption that makes possible the kind of information transfer I've called metalinguistic.

Bobby's assumption of the metalinguistic maxim is in a very clear sense one aspect of his more general assumption that Alphie is a cooperative conversational partner, an observation that suggests a deep continuity between metalinguistically conveyed information and more familiar modes of implicature. I won't have much to say on the question of whether metalinguistically conveyed information is in fact a variety of implicature in any deep sense. I suspect that the question itself may be terminological, though I won't be arguing for that either. Instead, what I'll do is observe a few more ways in which metalinguistically conveyed information is similar to implicatures, and then draw out a substantive philosophical point from those similarities.

First, sharpenings, like implicatures, can be cancelled. Suppose, for example, that Bobby wonders whether *athlete* is best used to refer to sporting humans only or to all sporting animals. To discover this, he chooses an animal, Secretariat, the physical characteristics of which are mutual knowledge between himself and Alphie, and asks Alphie whether Secretariat is an athlete. Alphie responds with (74).

> (74)   Well, if by *athlete*, you mean any animal that engages in sport, then yes, Secretariat is definitely an athlete.

By stipulating what his usage of *athlete* is to mean in this context, Alphie is able to express the literally true proposition that Secretariat is an animal that engages in sport without thereby committing himself to the claim that *athlete* is best used in that sense. Because that implication, carried metalinguistically, is the vehicle by which Bobby would have discovered how *athlete* is to most appropriately used, Alphie cannot learn what he aimed to learn from Bobby's utterance.

Similar cases show that, as with implicature, canceling needn't simply nullify the extra information; it can also directly contradict what ordinarily would be implied. Just as a speaker can directly and explicitly contradict a proposition that would normally be implicated:

(75)   You bet she's smart—in fact she's a genius!

so too can a speaker directly and explicitly contradict metalinguistic information that would ordinarily be carried by an utterance.

(76)   According to Dick Cheney's definition, waterboarding isn't torture—but
         that definition is rejected by legal experts.

Again, the speaker succeeds in expressing one true proposition—in this case the proposition that waterboarding fails to produce pain rising to the level of death, organ failure, or the permanent impairment of a significant body function—while at the same time explicitly denying another proposition that ordinarily would be implied by his choice of expression—namely, that Cheney's is an appropriate definition of *torture.*

So, like implicatures, metalinguistic inferences are made possible by a defeasible assumption of cooperativity on the part of the speaker. And, like implicatures, sharpenings can be explicitly cancelled by a speaker looking to limit or alter their commitments in making an utterance. The similarities may go even further however. There are, I believe, certain locutions that are most plausibly analyzed as outright floutings of the metalinguistic maxim. Those locutions involve speakers stating, sometimes for comedic or rhetorical effect, that they mean something wildly non-standard by their words. However, within locutions of that kind, there is an important distinction. Consider first the following passage from a review of the movie *Even Cowgirls Get the Blues* by Nathan Rabin.

(77)   I was a college freshman, appropriately enough, when I saw *Even Cowgirls Get The Blues* on videotape in 1994. It was a mind-blowing, revelatory experience. By "mind-blowing" I mean "forgettable," and by "revelatory" I mean "tedious."[58]

In this kind of statement, words are claimed by their own speaker to mean the opposite of what they usually do. Since Rabin, however inappropriate his usage of *mind-blowing* or *revelatory* may be, is not being deliberately uncooperative, this may seem like a plausible candidate for flouting. However, I don't believe that this is the proper analysis of the passage. The reason is that a much simpler analysis is available, namely sarcasm. Using one word to convey a meaning that ordinarily is the opposite is hardly an exotic rhetorical device. Whether Rabin's subsequent self-translation is literally true depends of course on what the English word *mean* means. But that won't matter for our purposes. What matters is that in this case there is a well-known pattern of usage according to which speakers can utter something literally false in order communicate the opposite. Rabin's utterance is a clear enough instance of that pattern that it is unnecessary to construct a more complex analysis.

Sarcasm however cannot be the correct analysis of superficially similar cases. Consider the following note, only a slight alteration of a real invitation.

(78)   Molly and I would like to see *The Da Vinci Code* tonight and wondered if you and Victoria are interested. And by "you and Victoria" I mean you and by "Molly and I" I mean Molly. I'm going to go but I'm not happy about it.

Let's call the invitation's author Josh, and its recipient Ted. Josh writes against the following background: he knows that he and Victoria will not like the movie, but he also knows that Molly and Ted will like the movie. Having made plans as a couple to see the movie, Josh and Molly would both, in a very ordinary sense, *like* to go see it. Likewise, assuming that Ted may succeed in persuading Victoria to see the movie, there is a reasonable hope that both Victoria and Ted would, as a couple, be *interested* in seeing it. The first sentence in the passage then, on by far the most natural interpretation, is very likely true. But if it is best analyzed as being true, then it is hardly a good candidate for

---

[58] Rabin (2009).

sarcasm. Furthermore, the glosses that Josh suggests in his next sentence are in no sense the *opposites* of what the targeted expressions would ordinarily mean. So when Josh proposes his non-standard gloss, it cannot be an ordinary case of sarcasm. Something more complex is going on.

On readings of *like* and *interested* that apply collectively to the respective couples —to Josh and Molly as a couple, and to Ted and Victoria as a couple—the first sentence in (78) is true. But on readings of *like* and *interested* that apply distributively to each of the individuals named, the first sentence is false. False, that is, unless by *Molly and I*, Josh meant Molly, and by *you and Victoria*, he meant Ted. On that highly unusual reading, the first sentence would be true—true in virtue of being entirely insensitive to the interests of Josh and of Victoria. My suggestion is that by proposing this unusual gloss on his own usage, Josh means to draw attention to just that tension between the two readings. Yes, as a couple, he and Molly would like to see *The Da Vinci* code. But lest the reader assume that this instance of *like* is *also* interpretable as distributing over Molly and him individually, Josh clarifies that to him, a reading of *like* according to which he would *like* to see *The Da Vinci Code* is about as natural as a reading of *Molly and I* that refers only to Molly.

As in Rabin's movie review, the question of whether Josh's translation claim is *true* depends on his intentions in uttering the first sentence and, crucially, on the meaning of the English word *mean*. On that subject, there is little reason to think that the English word *mean* picks out what philosophers and linguists pick out with their phrase *literally express*, so it is not implausible that Josh's claim about his own meaning is accurate. Even if the claim is false however, his expression could still be interpreted as a flouting of the metalinguistic maxim.[59] Josh uses (or claims to use) his words in a radically inappropriate (because massively non-standard) way. In that sense he appears to be acting uncooperatively. However, he does so in order to communicate something else

---

[59] That the proposition literally expressed is false would not prevent a speaker from successfully flouting a maxim to cooperative effect in standard cases of implicature. If the professor *falsely* claims that the student has good handwriting, that does not prevent her from thereby communicating that the student is poor in more relevant respects.

that more generally represents a cooperative (because clarifying) contribution to the discourse. That's flouting.

So the metalinguistic maxim, like the standard Gricean maxims, generates implications that go beyond the proposition literally expressed in an utterance. Just like the standard Gricean maxims, it can be flouted for cooperative communicative effect by a resourceful speaker. The implications generated by the metalinguistic maxim—sharpenings—can be cancelled just like implicatures. And, as we saw early, disagreements about those implications pattern with implicatures in their licensing of metalinguistic denials.

There is an important philosophical lesson to be drawn from this deep continuity between ordinary implicatures and sharpenings[60]. To see it, consider the relation between Gricean implicature and truth. To take one well known example, quantity implicatures allow us to give a semantic theory of numerical quantifiers according to which they literally establish only a lower bound. So the claim in (79) for example literally means only that John has *at least* five children[61].

(79)   John has five children.

The fact that (79) can almost always be used to communicate the proposition that John has exactly five children is explained by the familiar pragmatic story. One upshot of adopting this kind of story as opposed to a theory on which *five* means *exactly five* is that there is more truth telling. After all, supposing that John does have exactly five children, it's not only true to say that he has five children, but also that he has four children, or that he has three children, or that he has two children, or that he has one child. Indeed, in a debate between defenders of any two of those five expressions, neither speaker says anything false.

As philosophers, concerned to avoid the most pernicious forms of relativism, shouldn't this worry us? Obviously, it should not. The reason is that, when it comes to implicature, we know that a speaker communicates information beyond the proposition

---

[60] See Chapter 3 of this thesis for an elaboration of the argument below.

[61] See Horn (2004) for discussion.

they literally express. If that further information is wrong, then the speaker has said something defective, indeed something *mistaken* in a very clear sense, irrespective of the truth of the proposition she expressed. Thus, while it is important to establish that claims like the following are defective in *some* sense,

(80)   There are 49 states in the U.S.

(81)   Kant wrote two books.

(82)   There is one proton in the nucleus of a helium atom.[62]

there is nothing to fear from the truth of the propositions they express. A speaker uttering those claims speaks truly, but is mistaken in virtue of implicating something false, and is therefore criticizable in all the usual ways. As for a disagreements about these matters— one proton or two?—they will involve two speakers both expressing true propositions, but they are by no means unsubstantive. At most one of the speakers accurately implicates that there are exactly as many protons as in fact there are. That, clearly, is enough to avoid pernicious relativism.

Although the facts are less familiar, the situation is no different in the case of sharpening. This fact, unfortunately, can be obscured by the fact that in some situations it is not entirely clear what underwrites a word's being *appropriate*. Recall the debate about Secretariat:

(83)   Secretariat is an athlete.

(84)   Secretariat isn't an athlete.

We've stipulated that the speakers of (83) and (84) use *athlete* differently, and that their individual meanings are such as to render the two propositions they express consistent. The question is whether *athlete* can be used in such a way that it can refer to horses. Simplifying a bit, the debate is about the truth of the following propositions, communicated metalinguistically:

(85)   A usage of *athlete* according to which it does refer to horses is most
        appropriate to the conversation.

(86)   A usage of *athlete* according to which it does not refer to horses is most
        appropriate to the conversation.

---

[62] There are two protons in the nucleus of a helium atom.

Those propositions do conflict with each other. But which of them is right? Well, it's hard to say. It is not entirely clear what should go into deciding whether one or another usage of the word *athlete* is more appropriate to a given set of circumstances. If most metalinguistic disagreements are like this, then perhaps they are not so similar to the clear cut cases of true-but-mistaken utterances considered above.

Other cases of metalinguistic disagreement are more clear cut however. Recall the debate about Pluto, but this time, let's suppose that John is not a stubborn man on the street, but rather a student of astronomy, indeed a student of Betty's. The conversation to which *appropriateness* is relative is thus a scientific conversation, and the standards of appropriateness will be the standards of science. We'll maintain from the earlier scenario the fact that John uses *planet* in the ordinary English sense of the word. So now imagine that Betty has asked her students to name some planets and John volunteers the claim in (87):

(87)   Well, Pluto is a planet.

And Betty responds, as usual, with the claim in (88).

(88)   Actually, Pluto is not a planet; it's just a large asteroid.

As in the *athlete* case, we've stipulated that the two speakers here use their words differently, such that the propositions they express are mutually consistent. But in this case, that analysis may at first appear more philosophically menacing. After all, the disagreement between Betty and John is *verbal*; they simply mean different things by *planet*. Worse still, if John *does* accept Betty's correction, what looks for all the world like the learning of a scientific fact will reduce to mere terminology shift. It may begin to appear that an account positing a difference in meaning will commit us to pernicious relativism after all.

As should be clear by now, however, it commits us to nothing of the sort. Consider the propositions that, on this analysis, are metalinguistically conveyed by John and Betty's utterances, respectively.

(89)   A usage of *planet* according to which it can refer to Pluto is most
          appropriate to the conversation.

(90)   A usage of *planet* according to which it cannot refer to Pluto is most appropriate to the conversation.

Those propositions do conflict. Furthermore, the conflict between them is adjudicable. Recall that the conversation between John and Betty takes place in an astronomy classroom, i.e. a scientific context. In a context like that, words like *planet* are useful to the extent that the categories they carve out line up with interesting features of the world. In other words, a usage of *planet* (or, if you prefer, a concept of PLANET), is appropriate insofar as it carves the world at the joints. In a scientific context then, the appropriateness of a given usage is very much a matter of objective fact. The concept that Betty advocates via her usage does an objectively better job of categorizing the objects in the universe in a theoretically useful way than the concept that John advocates via his usage. If John is smart enough to change his usage on the basis of Betty's correction, that change will not be mere terminological shift, it will be terminological improvement.

The continuity between implicature and metalinguistic sharpening is thus crucial in opening up to philosophers a view of semantic theorizing that is often and unfairly considered philosophically unpalatable. A semantic theory that ascribes no meaning-stabilizing role to factors like deference, expertise[63] or metaphysical naturalness[64], one that hews closely to individual usage, and which renders many of our disagreements verbal will not bring with it the specter of pernicious relativism. Speakers constantly convey messages not just about the objects denoted by their terms, but also about the appropriateness of the concepts they employ. And, in just the contexts where it is important that disagreements be substantive, the appropriateness of concepts is very much a matter of objective reality. The very same factors that, on an externalist analysis, make it the case that John speaks falsely can, on this analysis, be mobilized to demonstrate that he is *mistaken* in virtue of employing a scientifically defective concept.

---

[63] As in Putnam (1975).

[64] As in Lewis (1983), or Weatherson (2003).

And as we saw with true claims like *there is one proton in the nucleus of a helium atom*, this notion of mistake is entirely adequate to the task of holding relativism at bay. A disagreement between two chemists about whether there is one proton or two in the nucleus of a helium atom does not lose substance simply because, on the most plausible semantic theory, both parties speak truly. The speakers communicate conflicting propositions, and the question of which of those propositions is true concerns the chemical nature of the universe. The fact that the communication of those propositions does not occur via literal expression is entirely irrelevant.

Precisely the same story can be told about sharpening uses. A disagreement between John and Betty about Pluto does not lack substance simply because each speaker literally expresses a true proposition. In advocating for alternative concepts of *planet,* the speakers take opposing stands on the appropriateness of their favored concept to a setting in which metaphysically natural categorization is a fundamental goal. And the fact that Betty's concept is *better* relative to that goal is no mere terminological stipulation.

The upshot is that theories of meaning that posit large amounts of individual variation, that do not rely on external factors as meaning-stabilizing forces, that hew very closely to speaker usage in their accounts of semantic content need not scare off philosophers of language. Such theories posit more truth telling than their competitors, just as the now-favored lower-bound analysis of numerical quantifiers posits more truth telling than its competitors. But in each case, the greater prevalence of truth telling poses no philosophical threat. A more nuanced view of communication—one that gives adequate weight to other channels along which information is conveyed—ensures that truth, while it may be easier, is not over-easy.

Finally, it should be noted that the particular similarities I have observed between implicatures and sharpenings are ultimately irrelevant to the philosophical point. Philosophers are by now very comfortable with the idea of a true-but-mistaken utterance when it comes to implicatures. That comfort level provides a good starting point when thinking about no-falsehood metalinguistic disagreements that turn on objective matters of fact. But the point would apply even if metalinguistic usages conveyed the

information they do via some radically unfamiliar mechanism. The important fact is that the information is conveyed, the concepts, parameter settings, etc. are advocated, and that some of those choices will be objectively better relative to the circumstances than others. The possibility of substantive disagreement in a discourse in which only true propositions are expressed follows naturally from any view of human communication sufficiently nuanced to capture those facts.

Chapter 3

Usage, Error, and Naturalness

1. Usage and Stability

Some disputes concern matters of fact and some disputes are verbal, so the story goes, but it is not always easy to tell the difference. When the question arises for philosophers, it is often when some intuitively substantive debate is threatened with the label *merely verbal*. The dispute in question may be non-philosophical, but of philosophical interest—are whales fish? is this painting beautiful?—or it may itself be a dispute among philosophers—do temporal parts exist? does an action's outcome matter to its being good?

In each case, the pattern is the same. The parties to some debate use their words in systematically differing ways. Naively, perhaps, that systematicity is taken as evidence that the parties to the debate employ different concepts, or use distinct but homophonous lexical items, or employ distinct polysenym for a word with a range of meanings, or set the parameters of contextually variable terminology differently, etc. By one mechanism or another, in other words, the speakers *don't mean the same thing by their words*. I will call this argument schema—the inference from systematically differing usage to some difference in meaning, whatever the mechanism—the *usage argument*.

In opposition to the usage argument, the *stability argument* goes as follows: If the speakers mean different things by their words, then their dispute is merely verbal; they are talking past one another. But the dispute is not verbal; on the contrary, it is substantive. By modus tollens, the parties to the debate must mean the same thing by

their words. Whatever the variation in usage, and whatever the systematicity of that variation, the meaning of the terminology in question must be stable.[65]

While this back and forth between usage and stability arises across subdisciplinary boundaries, the perspective taken on the issues is consistently linguistic. The arguments may of course be mobilized for philosophical reasons: The philosopher offering an instance of the usage argument, for example, may suspect that a particular debate is philosophically superficial. Or, as is often the case, the philosopher offering an instance of the stability argument may do so to defend a debate's status as empirically or philosophically substantive. Nevertheless, and whatever the domain in which the question arises, the consequences of a sound instance of either argument are semantic. The usage argument attempts to establish the semantic thesis that the meanings of the relevant terminology vary across parties to the debate.[66] Likewise, the the stability argument attempts to establish the semantic thesis that the meanings of the relevant terminology are stable across parties to the debate.

There is a difference however. The usage argument, whatever its flaws, is a rather straightforward semantic argument from empirical data to an empirical conclusion. The stability argument by contrast, begins with a philosophical premise—an assumption about the metaphysical status of the object of debate—but, like the usage argument, ends with an empirical conclusion about meaning. The stability argument, whatever its strengths, therefore creates a new burden—a burden that falls squarely on the shoulders of the philosopher of language and the semanticist: Whatever the facts of usage may be, other factors must be found to support a single-meaning theory. More generally, meaning, whatever it is, had better be the kind of thing that can be shared across scientific

---

[65] This simple argument is of course not the only argument on offer, in any particular domain, against a different-meanings view. But it is very, very widely applied, even in its most simplistic form. For the argument as applied to fish, see Lewis (1985) or Weatherson (2003). For the argument as applied to beauty, see MacFarlane (2007) or Lasersohn (2005). For the argument as applied to existence, see Sider (2009). For the argument as applied to goodness, see Smith (1994), Ch. 2.

[66] Or that parties to the debate set contextual parameters differently, or employ distinct poloysenyms, etc. The choice of a mechanism by a proponent of some instance of the usage argument depends on the empirical facts characterizing the specific terminology.

theories and across aesthetic perspectives; across ontological views and across longstanding, deeply entrenched ethical debates.

The question raised by this metaphilosophical pattern—by this repeated clash between the empirical import of usage and the philosophical need for stability—is thus largely methodological. What are the the criteria against which a theory of meaning is to be judged? How substantial a distance can be allowed between speaker usage and semantic theory? If the stability argument is generally successful, the burden imposed on the philosopher of language is that meaning must be the kind of thing that can be shared by two speakers who vary systematically in their usage and who, under the most careful and controlled of circumstances, will persist in their disagreements about the proper application of their words.

I will argue that this burden is too great to bear. No theory of meaning can plausibly do the things that are required by the stability argument. Furthermore, I will argue that across a wide range of cases, the inference from the systematicity of variant usage to differences in meaning is *not* naive—that it is precisely the correct inference to draw.[67] In other words, I will argue that across a wide range of cases, instances of the usage argument are sound. How then to respond to the stability argument? That argument, recall, relies on a certain inference: If the parties to a dispute mean different things by their words, then their dispute is merely verbal. If I am to reject the modus tollens of the stability argument, I must choose between accepting the modus ponens or rejecting that inference.

Accepting the modus ponens brings with it serious philosophical costs. After all, if, in a wide range of cases, our disputes involve speakers who do not mean the same things by their words, and if it follows that the disputes themselves are mere *talkings past*, then a most pernicious sort of relativism follows. Disagreements that, intuitively, represent substantive debate about objective matters of fact will be analyzed as "no-

_____

[67] Recall that there is a variety of mechanisms by which differences in meaning can be explained. Contextual variation, domain restriction, polysemy, as well as full-blown homophony are all options available to the semanticist. The usage argument does not replace the careful construction and defense of a specific semantic theory in a given domain, and the usage argument by itself can, at best, support the disjunction of those options.

fault," since, according to the conclusions of the usage argument, both parties will speak truly according to their own meanings. Disputes that are settled via substantive scientific progress will devolve into mere terminology shift.

Those consequences are indeed unpalatable. There may be some arguments that appear at first substantive but prove on careful analysis to be merely terminological. But an analysis concluding that all or almost all of our disputes are nothing more than shoving matches over terminology, freed from any objective mooring, is subject to the most trivial of reductios.

My job then in defending the usage argument will be to show how the inference from different meanings to *mere* talkings past can be rejected. To be clear, that job entails not *merely* showing that speakers can perceive themselves to be engaged in a real disagreement in cases where they mean different things by their words.[68] Rather, it will be to establish some way in which disagreements involving disparate meanings can be *adjudicated*. How, in a situation where there is intuitively some fact of the matter, could it be the case that at most one speaker is correct, and that change is not mere change but progress, even where both speakers speak truly? I think that this job can be done. Furthermore, I think it can done using the same resources that other philosophers have mobilized towards similar goals, but without incurring the costs they do. Before getting to that however, it will be worth a brief consideration of the logical space that these competing theories will occupy.

2. Everything is correct vs. Nothing is relevant

In a 2005 post for the blog Language Log[69], the linguist Geoffrey Pullum responds to an angry reader. The reader's complaint concerns the question of whether the linguists at Language Log are *descriptivists* or *prescriptivists*. The linguists of course claim to be descriptivists. Nevertheless, in an earlier post, Pullum had displayed the

_____

[68] I defend that weaker claim in Chapter 1 of this thesis.

[69] Pullum (January 25, 2005).

audacity to brand a naturally occurring sentence, taken from a newspaper article and involving a failure of subject/verb agreement, as ungrammatical. The commenter, Zink, complains that if Pullum genuinely is a descriptivist, and if someone "honestly makes a sentence," then Pullum must treat that sentence as a genuine datum, an "honest sentence in the language that actually is". Pullum responds

> [Z]ink cannot see any possibility of a position other than two extremes: on the left, that all honest efforts at uttering sentences are ipso facto correct; and on the right, that rules of grammar have an authority that derives from something independent of what any users of the language actually do. But there had better be a third position, because these two extreme ones are both utterly insane.

At one end of the spectrum then is the position Pullum labels *Everything Is Correct*. At the other, *Nothing Is Relevant*. Positions on the spectrum are differentiated by their answers to the question: How much actual speaker usage must be captured by the theory of grammar? If the answer is *Everything*, then there will be no relevant notion of speaker error. Any utterance, even those that most transparently involve the influence of what Chomsky calls *performance error*—processing difficulty, stuttering, or the most unexotic kind of misspeaking—must be generable by the theory of grammar.

At the other end of the spectrum is the view that the theory of grammar is accountable to *nothing* in the way of actual speaker usage. Whatever speakers say, whatever sounds acceptable to them, no matter how systematic or consistent their usage under the most controlled of circumstances, the theory of grammar may—accurately—describe them as wrong. At one end of the spectrum, then, there is altogether too much material to cover; the theory of grammar must also be a theory of processing and discourse and speech pathology and everything else. At the other end of the spectrum, there are no data at all. The theory of grammar can say whatever it likes and nothing from the realm of actual linguistic behavior can serve to constrain it.

Pullum's point is that any reasonable theory of grammar must plant its flag somewhere between the two extremes, somewhere where some defective instances of real world speech can be tossed out, but where there is still a robust set of facts constraining the theory which is, after all, empirical. A further point can be made about that choice of a location on the spectrum: It must be motivated. There is a lot of space between the two

extremes. If there is no *independently motivated* theory governing the specific distinctions that are made between performance errors and usages or intuitions that accurately reflect the grammatical competence, theorizing will be entirely unconstrained. Linguists could pick and choose those data that support their pet theory, while dismissing countervailing evidence as performance error.

Fortunately, in the domains Pullum is addressing and indeed across the subfields of linguistics, there is in general a clear enough idea of the other systems involved that this methodological free-for-all is avoided. Independently motivated theories of language processing, for example, can be drawn on to explain away usages or intuitions that seem to contradict otherwise well motivated grammatical theories, an example I'll return to below. Naturally, a finished and perfect theory of processing is not required to make progress in syntax. But as the theories advance in parallel, each must be able to draw on our best ideas about the other in selecting which data are most significant.

The spectrum Pullum describes does not just apply to theories of syntax. In the realm of semantic theory, there is also a need to allow for the possibility of speaker error without thereby untethering the theory from the empirical facts of language use. Because of the special role of the notion of truth in semantics, however, the choice of a location on the spectrum takes on greater philosophical significance. The question of which data are to constrain the theory is not a simple matter of usage or judgements of syntactic well-formedness.[70] Rather, it involves (among other things of course) speakers' first order intuitions about the world. In this context the right end of the spectrum—where any time a speaker "honestly makes a sentence" they *ipso facto* speak correctly—is not insane merely in virtue of making life implausibly difficult for the linguist. When the domain is semantics, and the conclusion is that every sincere utterance is correct, then we have concluded that everybody (provided they say what they really believe) speaks truly all of the time. In other words, the right end of the spectrum amounts to pernicious relativism.

---

[70] I believe, following Ludlow, that intuitions of syntactic well-formedness are better thought of as judgements of acceptability, *grammaticality* being a theoretical term, and thus not the kind of thing ordinary speakers have intuitions about. The subtleties surrounding this question won't matter for purposes of the discussion here. See Ludlow (manuscript), Chapter 2 for an extensive discussion.

Fortunately, the most extreme versions of this view can be dispensed with straight away. To see this, consider some simple examples. When a semanticist—let's say a field linguist—begins to construct a lexicon for a language she studies, she must of course base her work on the ways in which her informants actually use their words. But she will have to make certain assumption in her theorizing about which kinds of data are most important. If, for example, an informant points to a cat and describes it with a word the linguist has come to believe means *raccoon*, the linguist must make a calculation of a certain kind. Perhaps the community groups cats in with raccoons. Or perhaps the informant simply made a mistake. If the informant reports the judgement at dusk and from a great distance, and under other circumstances refuses to employ that description in reference to raccoons, then our linguist is well justified in writing off the potentially countervailing datum as a simple mistake on the informant's part.

Everyone then should be able to agree that there are some factors that must intervene between usage and a sensible semantic theory. Some of those factors will involve simple, one-off mistakes of the kind our linguist encounters above. Others will be psychological factors familiar from other subfields of linguistics where they play similar roles. Consider for example the expressions in (1) and (2).

(1)   The horse raced past the barn fell.

(2)   No head injury is too trivial to ignore.

The expression in (1) strikes many speakers of English as ungrammatical. But that judgement can be explained by the well-known difficulty of parsing so-called garden path sentences.[71] With enough explanation and reflection, many speakers come to see that the sentence is grammatical, provided it is understood that what fell was the horse, not the barn. The theory of grammar can therefore go ahead and describe the expression as well-formed, even if the raw usage data would seem to suggest otherwise.

Likewise, speakers tend to judge (2) as meaning that no head injury is trivial *enough* to ignore. Nevertheless, any standard semantic theory will predict that the proper interpretation of the sentence is roughly the opposite of that. But, as in the case of the

---

[71] See Pinker (2004), for a discussion of this example.

garden-path sentence, the damage from this seemingly false prediction is mitigated by our understanding of intervening factors. In this case, the reading that most speakers land on can be explained by the well established processing difficulty involved when negations, particularly negations that are lexicalized as in *ignore*, pile up in a single expression.[72] Again, the semantic theory can go ahead and predict the interpretation it does, even if that is at odds with actual speaker usage, because the coming apart of usage and theory is explained by an independently motivated generalization from another domain.

In addition to such one-off mistakes and psychological performance errors, there are other factors, specific to semantics, that are familiar from both linguistics and the philosophy of language. First, there are clear instances of rhetorical devices like sarcasm: When a speaker, with distinctive intonation, says that some transparently bad idea is *great*, we do not thereby replace our theory of *great* with one on which it can literally mean *bad*. Second, there are cases where speakers engage in games of pretend. The semantics of words that are used in the context of fiction or pretend is of course its own highly controversial issue, one that I won't engage here.[73] For now I will be content to grant that in the clearest cases of pretend—when a child refers to a bicycle as a *spaceship* for example—the linguist is free to disregard data that on the most superficial analysis would contradict an otherwise well supported theory.

Finally, and perhaps most relevantly, there are the ubiquitous cases of implicature, where a semantic theory is mobilized only in concert with a pragmatic theory to explain our usage of certain expressions. To take one well known example, our usage of numerical quantifiers exhibits patterns that are not captured by the now widely accepted semantic theory on its own[74]. According to that theory, numerical quantifiers establish a lower, but not an upper bound. Thus a claim like that in (3)

> (3)   John has five kids.

---

[72] See Wason and Reich (1979) or von Fintel (2004) for discussion.

[73] See Ludlow (2006) for an argument that in fact fictional contexts do not require non-literal interpretations of words.

[74] See Horn (2004) for discussion.

means literally that John has at least five kids. Nevertheless speakers typically use an expression like that in (3) to communicate the claim that John has *exactly* five kids. What comes between this pattern of usage and the semantic theory to explain the divergence is the familiar pragmatic story according to which speakers assume of their interlocutors that they will say the logically strongest relevant thing that they are in a position to say. In a typical context then, it is reasonable to assume that if John had more than five kids, the speaker of (3) would have said so.

Standing between speaker usage and semantic theory then—establishing some distance between the *everything is correct* end of the spectrum and a non-insane theory of meaning—are the usual suspects from other areas of linguistics: language processing, speech pathologies, etc. In addition, there are the familiar factors from semantics and pragmatics—conversational and conventional implicature, along with rhetorical devices like sarcasm. Finally, there are the inevitable cases of one-off mistakes: the speaker who calls a friend by the wrong name or calls a cat a raccoon in low light. Assuming that in the day to day work of the semanticist, factors like processing, sarcasm, and one-off errors will play a fairly marginal role[75], the idea so far is that the meaning of speakers words' is best described by a theory that accurately characterizes actual patterns of usage, once the influence of pragmatics has been factored out. Let's call this way of approaching semantics the Meaning is Use Minus Pragmatics Strategy, or MUMPS[76]. The question for us now is whether MUMPS establishes a *sufficient* distance from the right end of the spectrum to avoid the most alarming conclusions found there.

3. Not Enough Error

---

[75] This assumption could be wrong. If it is, it will not affect my arguments here, except, perhaps, for the terminology I employ.

[76] I should be clear that when I describe the view as one on which "meaning is use," I don't mean to imply that there are no facts of the matter about meaning aside from the facts of usage. I am not advocating anything like a strict "use theory" in the sense, for example, of Horwich (2004). I am merely describing a view according to which the methodology best suited to the task of constructing an accurate theory of meaning—whatever its metaphysical status—is one that closely follows the contours of actual patterns of usage. (Minus, in this case, the influence of pragmatics.)

Unfortunately, the answer might appear to be no. To see this, consider a field linguist working among English speakers of the 19th century[77]. The linguist discovers among many other things that her informants have a word, pronounced [fɪʃ], that they employ to describe animals that swim in the sea. In making her hypothesis more precise, our linguist questions her informants about a range of sea-dwelling animals and discovers that while they apply their word happily to trout and salmon, they refrain from applying it to otters and walruses. Finally, of course, she observes that her informants enthusiastically employ their word to describe whales. They do so sincerely and consistently, under the best of observational conditions and across a range of contexts varied enough to rule out any independently observed pragmatic influence. They are not being sarcastic and they are not playing pretend. Our linguist, having ruled out each of the factors that, according to MUMPS, can come between real patterns of usage and proper theory, determines that the (19th century) English word [fɪʃ] (i.e. *fish)* picks out a category of sea-dwelling animals that includes both fish and whales.

Naturally, this conclusion may be considered cause for some alarm. After all, if it is right that 19th century English speakers mean *fish and whales* by their word *fish,* then their utterances of expressions like (4) were *true*.

(4)   Whales are fish.

But whales are not fish. They are mammals. 19th century English speakers were wrong when they said that whales are fish, and when they discovered that whales are mammals, that discovery constituted new knowledge about the nature of whales and fish, not a mere change in terminology. It seems then that despite its accounting for pragmatic interference, performance errors, and one-off mistakes, the methodology described in MUMPS still does not have the resources to hold relativism at bay. The reason, it would appear, is that although it allows for one-off mistakes, it does not allow for the kind of mistake we ascribe to 19th century English speakers, namely systematic, pervasive error. Obviously, there is nothing terribly unique about the case of whales and fish, so if the

_____

[77] Obviously the factual accuracy of this story about earlier versions of English is not my highest priority here.

linguist is right to apply MUMPS in this case, then she will be right to apply it in a whole range of cases that will bring with them relativism about our most intuitively substantive debates and discoveries.

That's a pretty bad conclusion. Naturally enough, then, philosophers have worked to avoid it. The problem they have diagnosed (in other terminology) is that MUMPS appears to link semantic theory, and thus our understanding of what it takes to say true things, too closely to the ways in which speakers happen to talk. If the problem is that there is not enough stuff intervening between actual patterns of usage and semantic theory, then the solution is to add more stuff. In other words, the solution is to further decouple a theory of meaning from speaker usage, and in decoupling it from usage, to link it to some other factor that can enforce stability of meaning across speakers who use their words in very different ways, and that can tether meaning to the worldab in ways that speakers can be entirely unaware of.

4. The Balance Theory

Weatherson (2003), expanding on a proposal from Lewis (1983) (himself taking up a suggestion from Merrill (1980)), describes just such a solution. Weatherson proposes that a theory of meaning must strike a sometimes tenuous balance between two competing factors: on the one hand it must capture many of the facts about actual usage, and on the other, it must posit as the meanings of our predicates properties that are *metaphysically natural.*

> [F]or any predicate t and property F, we want F to meet two requirements before we say it is the meaning of t. We want this meaning assignment to validate many of our pre-theoretic intuitions […] and we want F to be reasonably natural […]. In hard cases, these requirements pull in opposite directions; the meaning of t is the property which on balance does best.[78]

---

[78] Weatherson (2003), p. 9. On Merrill's original suggestion, naturalness is an all-or-nothing issue. Lewis jettisons that assumption, and constructs a theory on which naturalness, and thus eligibility, are matters of degree. (See Lewis (1984), p. 227.) Weatherson clearly follows Lewis in this regard. My arguments against Lewisian theories like Weatherson's will apply even given this more nuanced account of naturalness.

In other words, there is, in addition to performance errors, one-off mistakes, and the influence of pragmatics, another factor coming between usage and theory. And this factor is big. Our usage of our words, on Weatherson's view, can diverge widely and systematically from the usage that, given their meanings, would be appropriate. "The reason," he says, "is that highly natural properties are pretty thin on the ground; one's dispositions to use a term have to change quite a lot before they get into the orbit of a distinct natural property."[79] If my usage of some predicate comes anywhere close to picking out a certain metaphysically natural property, then that property is the best candidate for a theory of what I mean by that word, even if in a fairly wide range of borderline cases, I persist in diverging from that property in my actual application of the predicate.

Weatherson observes that on this account, it is possible to draw a clear distinction between disputes about meaning and disputes about fact.

> In some disputes, the same meaning postulate does best on balance at capturing the dispositions of each party. I say that here the parties mean the same thing by their words, and the dispute is a dispute about facts. In others, the difference will be so great that different meaning postulates do best at capturing the dispositions of the competing parties. In these cases, I say the dispute is a dispute about meaning.[80]

That same distance from usage that allows Weatherson to analyze some persistent disputes as disputes about fact allows him to make sense of speakers' being systematically mistaken in their application of their words. This, Weatherson argues, is a result that makes his view preferable to alternatives that link meanings more closely to actual usage.

> The hypothesis that when we alter intuitions because of a theory we always change meanings, on the other hand, is not even plausible. When the ancients said "Whales are fish," or "The sun is not a star," they simply said false sentences. That is, they said that whales are fish, and believed that the sun is not a star.[81]

---

[79] Weatherson (2003), p. 10.

[80] *Ibid*.

[81] *Ibid.* p. 11.

That result about whales and fish is exactly what MUMPS couldn't get, and so it seems that Weatherson has solved the problem we were faced with when performance errors, one-off mistakes, and pragmatics were all that stood between the theory and the relativistic right end of the spectrum. According to MUMPS, "the ancients" expressed true propositions when they said *whales are fish*, since they happily assented to expressions like *whales are fish* under circumstances where the influence of those other factors could be ruled out.

On Weatherson's story, by contrast, the ancients got close *enough* to picking out some scientifically respectable fishy category—something, perhaps, like what we moderns mean by *fish*—for that more respectable fishy category to be the meaning of their word. Supposing that their utterances of *whales are fish* expressed the proposition that whales are members of that more respectable category—something like the proposition that whales are fish—it is clear that indeed they said something false. Whales are not fish, after all. They are mammals.

5. Too Much Error

Weatherson's Lewisian balance theory of meaning avoids the alarming philosophical consequences that seem to follow from MUMPS. But that philosophical result comes at a steep empirical cost. The cost, like the problems he aims to avoid, is best understood with reference to the spectrum described above. So far, we have devoted most of our attention to left end of the spectrum, and thus the problems that have been most salient have been *everything is correct* type problems, namely, not enough error. But now let's turn briefly to the right end of the spectrum. What exactly is it that makes *nothing is relevant* insane? Well, first, there is the methodological nightmare that results from a total absence of data. If speaker usage is entirely irrelevant to a theory of meaning, then it is not clear how such a theory could even get started.

What, then, about slightly less extreme versions of far-right theories of meaning? To see what goes wrong with theories that allow for some data, but still reside to close to

the right end of the spectrum, consider the hypothetical example of extreme prescriptivism. What I'll mean by *extreme prescriptivism* is a theory that purports to describe the actual facts of a language (as opposed to, say, offering stylistic advice), but that in offering those descriptions, aims primarily to vindicate traditional grammatical rules of the kind that are taught to school children. What, for example, has gone wrong with a syntactic theory according to which any English expression that ends with a preposition is ungrammatical? Well, for one thing, such a theory would predict that the expressions in (4)-(6) are ungrammatical.

(4)   That was the person I was talking to.

(5)   I don't know which student I'm supposed to meet with.

(6)   He's always telling me about problems that I'm already aware of.

Those predictions appear to be wrong, since (4)-(6) seem for all the world to be grammatical.

Of course, on any theory not at the extreme left end of the spectrum, it is possible for speakers to mistaken in their usage. But mere possibility is not enough. As emphasized above, if speakers can be accused of mistaken usage with no *independent motivation*, the theory itself will be entirely unmotivated. In the case of processing errors, or one-off mistakes, or pragmatic interference, we can make sense of speakers' having intuitions at odds with their own linguistic competence. What, on the extreme prescriptivist view, would allow us to make sense of speakers across a wide range of contexts, in casual speech and in careful writing, systematically and sincerely and enthusiastically uttering expressions that end in prepositions? Short of some theoretical *deus ex machina* to explain that, the prescriptivist claim that grammatical English sentences cannot end with a preposition is unsupportable.

Extreme prescriptivism is a straw man. But its shortcomings demonstrate the problem with theories that veer too far from usage. That problem is *too much error*, and it is shared by Weatherson's non-straw-based balance theory. Ultimately, I will argue that Weatherson (and therefore Lewis and others) veers too far from usage even in his claims about the ancients and their whale-talk, but to see the problem that Lewis and Weatherson

have with error, consider first the more recent phenomenon of bird watching: Bird watchers care about the natural world. In their efforts to see as many species of birds as they can, they typically pay close attention to the science of ornithology and its ongoing efforts to categorize bird life in ways that are increasingly useful, natural, and theoretically well-motivated. As a result, it is often the case that birders must revise their lists of sighted birds, either to reduce them in response to species mergers, or to augment them, if both subcategories of a species split have been seen.

In some cases, however, ornithologists go a distinction too far for the purposes of recreational bird watching. As the science advances, ornithologists posit new species distinctions that can be drawn only on the basis of DNA analysis carried out in the lab.[82] In those cases, the discoveries they make are of categories that, while they are more metaphysically natural, cannot be properly distinguished on the basis of even the best observational evidence in the field. Nor is this a phenomenon entirely limited to recent genetic research. Similar cases involve earlier distinctions made on the basis of very close examinations of a dead specimen, not a mode of observation available to the typical birder.

In many of the cases where this happens, birders *knowingly cease to defer* to the ornithologists in their usage of species terminology. When the categories posited by the ornithologist are simply too fine grained to be useful for the practice of birdwatching, birders reveal their usage to be governed by a norm that does not strictly align itself with our best scientific theories. Rather, the norm governing usage of terms like *species,* as well as individual species names, is more like this: get as close as possible to scientific naturalness while remaining observationally distinguishable given the tools of birdwatching. As noted, people who take up birding typically care intensely about the objective facts of what they've seen. But beyond a certain point, they simply don't care what the scientists tell them about species. They care about what, in principle, can be seen with a good pair of binoculars. These people talk to each other a lot, and they say things sincerely and under good observational conditions that are judged consistently by

_____

82 See Price (2007) for an overview of the ornithological issues.

their peers to be true. But if the meanings of their terms were tethered exclusively to the categories that are be useful for the scientific study of bird life, rather than the recreational practice of birdwatching, vast swaths of their discourse would be false.

What the birding case demonstrates is that in some contexts we simply don't value metaphysical naturalness as much as other factors that influence the usefulness of our terms, and that the facts about what we value are relevant to any plausible theory of our meanings. The birding case shows that this can be true even in situations where we do care about the "joints" in the world, but where that consideration must be balanced against others. However, the extent to which we value metaphysical naturalness will vary dramatically, and in some contexts, we will value it even less than birders do, even as our usage remains highly patterned and systematic with respect to other factors. The Lewisian balance theory simply doesn't have the resources to give a nuanced account of these facts, and the result is that it saddles the semanticist studying birdwatchers with a massively implausible error theory. For a Lewisian like Weatherson, any usage that systematically but subtly diverges from a nearby metaphysically natural categorization of the world—*even in contexts where that divergence is knowing and intentional*—will be analyzed as systematic error. After all, the birder comes very close to the scientific concept of species in their usage, certainly close enough that there is no other metaphysically natural concept of species whose orbit their usage could fall into. The fact that metaphysical naturalness is not the kind of naturalness that the birder cares about should be important to anyone giving a theory of a word like *species* as it is used by birders. But it has no role to play on the balance theory.

The first serious problem with a Lewisian balance theory then is a problem characteristic of any theory that has moved too far to the right on the spectrum: too much error. In this case, the problem is that the methodology Weatherson describes commits him to pervasive error theory in any context that does not happen to be inhabited by scientists or philosophers. If you link meaning to metaphysical naturalness across the board, you can preserve the possibility of systematic error in scientific contexts (perhaps a benefit, insofar as it keeps relativism at bay), but at the steep cost of being saddled with

error theories across the vast range of contexts in which speakers simply do not care, or do not care exclusively, about metaphysical naturalness, and are happy to regulate their usage according to other norms.

There is another serious problem with tethering a theory of meaning to metaphysical naturalness. Building metaphysical naturalness into a theory of meaning leaves no role for what I will call *psychological naturalness*. To see this, consider some observations made in Chomsky (2007) about the range of things we can say regarding the Charles River.

> It would still be the Charles River, under some circumstances, if it were divided into separate streams that converged in some new place, or if any H_2O that happens to be in it were replaced by chemicals from an upstream manufacturing plant. On the other hand, under trivial changes it would no longer be a river at all: for example, directing it between fixed boundaries and using it for shipping freight (in which case it would be a canal, not a river), or hardening the surface to the glassy state by some near-undetectable physical change, painting a line down the middle, and using it to drive to Boston (in which case it would be a highway). And on to further intricacies, as we can easily determine.

Chomsky observes that our ordinary usage of terms like *river* is highly, almost shockingly, complex, and that the range of real world objects that those words can be used to refer to seems to change and shift in ways that have little to do with our theoretical understanding of the natural categories that comprise the world. From those observations, a range of conclusions could be drawn. One thing that I think follows clearly, however, is that a theory of the meanings of our words should make reference to a certain kind of naturalness—but the naturalness will be psychological, not metaphysical.

To see what I mean, consider another example of the complexity in our usage of a seemingly simple term. In this case, Chomsky riffs on the odd properties of *houses*.

> If I see the house, I see its exterior surface; seeing the interior surface does not suffice. […] But the house is not just its exterior surface, a geometrical entity. If Peter and Mary are equidistant from the surface—Peter inside and Mary outside—Peter is not near the house, but Mary might be, depending on the current conditions for nearness. The house can have chairs inside it or outside it, consistent with its being regarded as a surface. But while those outside may be near it, those inside are necessarily not. So the house involves its exterior surface and its interior. But the interior is abstractly conceived; it is the same house if I

fill it with cheese or move the walls—though if I clean the house I may interact only with things in the interior space, and I am referring only to these when I say that the house is a mess or needs to be redecorated.[83]

Etcetera. Chomsky observes, in other words, that "the internal conditions on meaning are rich, complex, and unsuspected, in fact barely known."[84] What we do know, however, is that those complex conditions are not random.[85] As Chomsky himself observes, the subtle patterns governing our usage of *house* appear to be consistent—even cross-linguistically—across a wide range of so-called "container" terms.[86] The conclusion to draw from such a pattern is not a wide-ranging skepticism about the reference relation or about the world-directed nature of our speech. Rather it is that we have been applying the wrong standards in thinking about the categories of object that we use our words to refer to.

The collection of things that we refer to using a word like *house* may form a rather bizarre category. It certainly is unlikely to carve the world at the kind of joint that a scientist (at least a scientist studying houses) might be interested in. But the features that make it naturalistically bizarre make it psychologically revealing. The fact that those strange features of our usage of *house* are shared by our usages of *box, igloo, airplane*, and *lean-to*, for example, (and especially the fact, if it proves true, that those features are shared cross-linguistically) has a great deal to tell us about the ways that people naturally carve up the world. I'll leave the task of saying just what we should learn from those patterns to the lexical semanticists, conceptual analysts, and cognitive psychologists. But the fact that those patterns exist tells us that there is something methodologically suspect about tethering a theory of meaning to a kind of naturalness that is disconnected from the idiosyncrasies of human cognition. The contours of our concepts are highly complex and

---

[83] Chomsky (2000), p. 36. These particular examples—regarding rivers and houses—may or may not persuade. Alternative examples are plentiful and quickly multiply. The important and more general point is that our use of even ordinary terms is patterned in subtle and sometimes surprising ways, and that those patterns do not cleanly line up with metaphysically natural categories of objects.

[84] *Ibid.*

[85] Thanks to Sam Epstein for helping me to clarify this point.

[86] Chomsky (2000). p. 35.

patterned, but the patterns that emerge when we look carefully have little to do with metaphysical naturalness. Applying metaphysical naturalness as a constraint on a theory of meaning therefore usurps the role that psychological naturalness should be playing, and hijacks the projects of lexical semantics and psychologically naturalistic conceptual analysis. Such a hijacking would perhaps be a worthwhile cost if the alternative, more heavily use-based theories did indeed have consequences as undesirable as pernicious relativism. Fortunately, they do not.

6. MUMPS

MUMPS, recall, is our name for a view on which the correct semantic theory for a speaker is a theory that accurately characterizes the way that speaker uses their words, once the influences of performance errors, one-off mistakes, and pragmatics are factored out.[87] MUMPS has I think a certain *prima facie* plausibility, and it has the advantage of offering a kind of methodological continuity between semantics and the other subfields of linguistics. In syntax, for example, the operating assumption is something very much like "grammaticality is use minus performance error," once *use* is conceived broadly enough to include judgements of acceptability both in normal speech and in the context of more focused linguistic thought experiments.[88] Similar assumptions are made regarding analogous notions like *possible word* in phonology.

If we think of semantics as a branch of empirical linguistics then, we should grant some benefit of the doubt to the idea that usage plays a role in semantics that is not massively dissimilar to the role that it plays in syntax or in phonology. But that assumption, plausible as it may be, is seriously undermined by the observation that according to MUMPS, the ancients spoke truly when they said things like *whales are fish.* The real danger here is that if the semanticist posits only as much error as the

---

[87] Other factors will of course play a role in more fine-grained theory selection, including a semantic theory's interaction with syntactic and phonological theory.

[88] See, for example, Chomsky (1965).

syntactician or the phonologist, then truth will be, as Lewis puts it, "over-easy".[89] Perhaps the methodological disconnect between semantics and other parts of linguistics is real. Perhaps we simply need more error in semantics than we do in syntax or phonology.

I think though that we do not. The reason is that the kind of error that the semanticist ascribes to speakers who, for example, make a one-off mistake, is very specific. It is an error that consists in the expression of a falsehood. But when we speak, we aim to accomplish much more than merely to express truths. Our aims when we speak are myriad, but they include, to return to a familiar example, the goal of communicating things that are as relevant and as informative as possible. We also aim to express our propositions with words and concepts that suit the goals of the particular conversational setting that we find ourselves in. When the full range of our communicative goals is taken into account, it becomes clear that, just as we aim to do more than say true things, we can err in virtue of more than saying false things. This observation is familiar in cases of conversational implicature. But when it is considered more broadly, it speaks directly to the cases that seem most alarming when MUMPS is taken seriously as a theory of meaning.

Consider briefly, then, those familiar cases of conversational implicature.

(7)   There is one proton in the nucleus of a helium atom.

The speaker of (7) is mistaken. There are two protons in the nucleus of a helium atom. Nevertheless, the undeniable fact that the speaker of (7) makes a mistake does not commit us to any particular view of the nature of that mistake. As a matter of fact, given the widely held assumptions about the semantics of numerical quantifiers referred to above, the speaker of (7) speaks truly. Nevertheless, that speaker is mistaken in virtue of *implicating* a false proposition, namely the proposition that there is *exactly* one proton in the nucleus of a helium atom.

Clearly then, the notion of *mistake* that is necessary to make sense of our scientific discourse is not so specific as to make particular demands of our semantic theory. This is equally true with respect to our selection of which concepts to employ in

---

[89] Lewis (1984), p. 225.

the course of inquiry. So consider precisely the case that seems most problematic for the proponent of MUMPS, the case of the ancients and their mistaken views about whales, and let us go ahead and stipulate what may seem to be a philosophical worst case scenario: the ancients really did mean something different by the word *fish* than we do. Our dispute with them is terminological, a disagreement about meaning. Their discovery that whales are mammals—their shift in theory—really did amount to a shift in terminology. The hypothesis that "when we alter intuitions because of a theory we always change meanings,"—the hypothesis that according to Weatherson is not even plausible—is, for the sake of argument, true.

In Chapter 1 of this thesis, I argue that the kind of disagreement that concerns the proper application of a word can be perceived by speakers as a robust, substantive disagreement like any other. And in Chapter 2, I argue that those disagreements, what I call *metalinguistic disagreements*, behave linguistically very much like the kind of disagreement over implicatures that might occur in a dispute over the number of protons in a helium nucleus. Empirically then, the hypothesis that a cross-theoretical dispute about the nature of fish is in fact a disagreement about meaning is not implausible at all. Such a terminological dispute might very easily be perceived by speakers as an ordinary, substantive dispute, and its effects on the discourse in which it is situated—its capacity or incapacity to license various kinds of denials for example—would be just as they in fact are. So when it comes to making the right kind of predictions about the data, the kind of theory that would result from an application of MUMPS seems entirely plausible. The problem with MUMPS is thus exclusively philosophical.

So far, however, that philosophical problem remains. After all, it is not enough that speakers *perceive* their disagreement to be substantive in this case. We must, if pernicious relativism is to be kept at bay, find some accounting of how this disagreement about meaning could *be* substantive. But that task is not as difficult as it may seem. On the hypothesis we are considering, the parties to the fish dispute each express true propositions of course. But that is not all they do. They also advocate distinct meanings for their word *fish*. In other words, though each may in principle agree with the

proposition expressed by the other, they disagree about which meaning of *fish* is better. That point of disagreement raises the obvious question: *is* one of their meanings better? Well, it seems that in some cases one of their meanings is better, and in others, it's a bit harder to say. In the context of a marine biology classroom, the modern meaning for *fish* is *undoubtedly* better. It is better not as a matter of convention, or social practice, or linguistic stipulation. It is objectively better, given the goals of marine biology, because the set of objects it categorizes go together more naturally with respect to the relevant biological considerations than the set of objects classed together by the ancient's word *fish*.

This dispute about meaning then—the dispute over which word *fish* is more appropriate, considered in the context of a biology classroom—*is* a dispute about fact. The modern speaker's choice of concept is objectively better than the ancient's in light of objective, discoverable facts about the natural world. In fact the very same facts that on the balance theory would make it the case that the ancient speaks falsely are the facts that, on this analysis, make it the case that the ancient speaker's concept is inferior to the modern speaker's. The only difference is that on this analysis, those facts do not play a meaning-determining role. The fact that in their dispute, both speakers happen to literally express true propositions has no more bearing on the question of whether their dispute is substantive than it would in the case of a dispute about the number of protons in the nucleus of a helium atom. The distinction that Weatherson aims to clarify is a false choice. Some disagreements about meaning are disagreements about fact.

What then about non-scientific contexts? Suppose that the dispute occurs, not in a marine biology classroom, but on a whaling ship. Here—as with the birdwatchers—what matters is not the goals of natural science, but rather the goals of one of the many other activities that humans engage in, in this case the goals of whaling and the day to day requirements of life on the sea. In this context, it is far less clear which meaning for *fish* is better. If anything, the balance shifts towards the ancient's usage, as the biological distinctions that justify our modern usage are of little use to the sailor.

Does this spell trouble for MUMPS? Of course not. Contrary to Weatherson's claim, it is not at all clear, much less platitudinous, that when the ancients (particularly the sailors among them) uttered an expression like (8), they were expressing the belief that whales are fish—not when the modern usage of *fish* is influenced by so many factors that mattered so little to the ancients. If the balance theory is correct, then the semanticist is saddled with an empirically unjustifiable error theory of the sailor's speech. On MUMPS, by contrast, once the full range of our communicative goals is taken into account, we can give an empirically defensible semantic theory for both the ancient sailor and the ancient biologist. Both spoke truly when they uttered the sentence *whales are fish*. That result is consistent with there being a clear and robust sense in which the ancient biologist was mistaken. Truth is easy on this account, but it does not follow that error is over-hard.

We've seen then that the amount of false-speaking that is possible given the perfectly ordinary factors intervening between usage and theory—performance errors, one-off mistakes, and pragmatics—is sufficient both to provide a sensible theory of meaning and to hold off the threat of relativism. The reason is that the category of error required to hold off the threat of relativism was never isomorphic with false-speaking. As philosophers have known since Grice, speakers can err in ways that go beyond the expression of false propositions. In the practice of an activity like scientific inquiry, some of those goals, including the discovery and employment of the concepts that cut the world close to the joints, are every bit as sensitive to objective fact as the goal of saying things that are true.

One final point remains however to fully defend a MUMPS view of meaning from its Lewisian competitor. Lewis, in proposing his original version of the balance theory, aims not only to prevent over easy truth, but also to resolve other puzzles about meaning, most notably Putnam's model theoretic argument, and Kripke's reading of Wittgensteinian skepticism about meaning. I will not have the space here to fully address either of those, but I'd like to suggest that the resources available to a defender of MUMPS view are at least as robust as those available on the balance theory.

Both "Kripkenstein" and the model theoretic argument raise problems for a theory of meaning by arguing that usage, no matter how closely a theory reflects it, is not by itself enough to determine which among some set of importantly non-identical interpretations is correct. For Putnam, the startling conclusion is that any theory, given certain plausible constraints, can be interpreted in such a way that all of the sentences comprising it are true. For Kripke, the problem is that any term can be interpreted in some implausible way that is indistinguishable from the intended interpretation given past usage. Both Putnam and Kripke's Wittgenstein were happy enough to bite the bullet with respect to the startling conclusions of their arguments. But more optimistic philosophers of language (like Lewis) have worked to avoid those conclusions.

For Lewis, the way to avoid the unpalatable conclusions reached by Putnam and Kripke is to fill in the gaps left by usage with universals, i.e. with the resources provided by the account of metaphysical naturalness. On his account, as on Weatherson's more recent version, some properties are themselves more natural than others, and that naturalness thereby makes them more "eligible" to be the properties that are picked out by our predicates. While our usage of a word like *plus*, then, may not in itself rule out quus as the correct interpretation, the very fact that the plus function is more natural in the relevant sense makes it, and not the unnatural quus function, the referent of our term[90]. Likewise for any *grue*-some predicate. If it competes with some more metaphysically natural property to be the meaning of our words, and if that more metaphysically natural property is roughly as good at making sense of much, if not all, of our actual usage, then the more natural property is the meaning of the term.

I think that similar considerations provide analogous resources for solutions on behalf of MUMPS. Take Kripke first. While it can be acknowledged that even the most closely use-based theory of our usage of the word *plus* will be inadequate to the task of ruling out quus-like alternatives, there is a clear sense in which plus is more *psychologically* natural than quus. (Or at least, a sense that is at least as clear as the sense

---

[90] This is how Lewis rejects Putnam's premise that any proposed constraint on reference is "just more theory". The naturalness-based eligibility of some property is entirely independent of our intentions to refer.

in which plus is *metaphysically* more natural than quus.) How to provide a non-question-begging analysis of psychological naturalness that brands quus less natural than plus? Well, it might be hard. But it seems unlikely to be radically more difficult than the project of providing a non-question-begging analysis of metaphysical naturalness that rules quus less natural than plus. In either case, the underdetermination gaps left by the patterns we exhibit in our usage can be filled by looking to the patterns that, from previous work on similar problems, we've grown to expect.

There will of course remain an element of underdetermination. But for that to be a special problem for MUMPS, we would have to be convinced that such underdetermination is in this case significantly more threatening or insidious than the underdetermination that is a fact of life in any scientific inquiry, or that attaches to any theory of meaning. The semanticist employing the methodology described in MUMPS has, in answer to the challenges of Kripke and Putnam, the very same resources that Lewis has, since, after all the semanticist can still describe a speaker as mistaken in virtue of failing to pick out the categories that are most appropriate given their conversational goals. And in addition to those resources, they have the further resources of psychological naturalness, acting as a background set of assumptions against which the partial evidence provided by usage can be judged. I find it very implausible to think that with the resources of both psychological and metaphysical naturalness at their disposal, the proponent of MUMPS will be *worse* off than the proponent of the balance theory in their responses to Kripke and Putnam.

7. Philosophical Upshot

To return then to the problem with which we began, we have seen that the recurring tension between the empirical demands of variant usage and the philosophical threat of relativism has caught philosophers in a false dilemma. When speakers use their words in systematically different ways, we are right to take that as evidence that they mean different things, whether that difference is captured in any particular instance by

contextual variation, polysemy, homophony, or whatever. The threat of relativism need not be addressed at all in the construction of such a theory, since the possibility of pervasive error does not depend on the possibility of pervasive falsehood. A semantic theory can ascribe meanings that hew closely to speaker usage, while the philosophical work of accounting for mistaken theories is taken care of by notions like betterness of concept that are underwritten by the very philosophical theories that Lewis and Weatherson needlessly load into reference determination.

The alternative view, on which the factors that make one meaning better than another are in part meaning-determinative, preserves the stability of content across speakers and across theories and across persistent debate. But it does so at a steep cost. That cost includes an unexplained methodological disconnect between semantics and other areas of linguistics, error theories where they are not plausible, and finally the failure to ascribe a properly significant role to psychologistic considerations in what, ultimately, is an empirical, psychological inquiry. These concerns arise for the balance theory even without our inquiring into the empirically mysterious mechanism by which the meanings of our lexical items are tethered to the largely undiscovered categories of the world. The view that I have called MUMPS thus has significant advantages over the balance theory. And, crucially, it does not have the philosophical drawbacks that heavily use-based theories are ordinarily taken to have. It does not entail pernicious relativism. And it is no worse off than the alternatives, and may in fact be better off, when it comes to the challenges of Kripke and Putnam.

I'd like to conclude then by reconsidering briefly and in a general way the examples with which I began the paper. I have already discussed whales and fish at length. And I discuss the question of whether some artwork is beautiful in Chapter 1. But what about pressing, unresolved, and substantive-seeming questions like the existence of temporal parts, or the relevance of an action's consequences to its goodness? Here, the main point I hope to make is that while the view I advocate has important consequences for those debates, particularly in opening up previously shunned areas in

move space, it does not do so much philosophical work as to render itself implausibly powerful.

Say, for example, that we adopt an analysis of the word *good* according to which consequentialists and deontologists denote distinct concepts with their terms.[91] Supposing we adopted such an analysis, we could now be clear on a number of things that may otherwise have been unclear. We would know, for example, that the important ethical project does not consist in trying to describe some concept of goodness that is both specific enough to do the philosophical work that such a concept must do and yet permissive enough that it could plausibly be the concept shared by both parties to the ethical debate. Such a project would be misguided on this view. On the other hand, we can also be clear about a number of independent but philosophically pressing alternative projects. For example, there is the work to be done in describing the concept that deontologists *do* employ. There is the project of describing the concept that consequentialists employ. Most importantly, on an analysis according to which the debate between the two turns on their advocating distinct concepts of the good, there is the question of which concept is better.

When the concept in question was *fish*, we knew roughly what to say. When the conversation about fish occurs in a biology classroom, it is very clear what betterness resides in, and it is also very clear which concept has more of it. When the conversation occurs on a whaling ship, it is perhaps less clear what betterness resides in, but then again, it is also less clear which of the disputants is right. Those results are about as good as one could hope for—the easy cases are still easy and the hard cases are still hard. But what about the concept of *good*? What does it mean for one concept of goodness to be better than another? The answer of course is that it is very hard to say. But as in the whaling case, that difficulty is the right result. The question of what would make a claim about the good *true* is very difficult to answer on a more traditional analysis. The fact that the parallel question is still difficult on my view is a point in its favor. The fact that the question, while it may still be hard, is a bit clearer, is a big point in its favor.

---

[91] This is not a view I intend to advocate here.

The same is true with respect to temporal parts. In that area in particular, there is an active meta-debate about what is at issue in the object-level debate.[92] The question at issue in the meta-debate is whether the object level debate about the temporal nature of material objects is substantive or terminological. Those, like Hirsch[93], who advocate the view that it is terminological, do so in part to vindicate their more basic view that the debate about temporal parts is defective. Those, like Sider[94], who advocate the view that it is not terminological, defend their position in part on the grounds that the debate must be substantive. One position that has received little attention, however, is the view that the debate is terminological but nevertheless substantive.

The ability to bring that position to the table is an important consequence of the view I advocate here. But if anything, it makes the metaphysical question harder. After all, we may, if we adopt the view in MUMPS, determine that the debate over temporal parts is terminological. But that will not tell us one way or the other whether it is substantive. The question of whether it is substantive must be answered by some accounting of what it would mean for one way of talking about material objects to be better than some other way. That question will not be answered by any theory, however nuanced, of the nature of communication.

To see the more general point, consider again the examples raised as concerns for the balance theory: the birdwatching-style constraints on a word like *species*, and the Chomskian internal constraints on a word like *house*. Both sets of considerations concern *naturalness*, but they are clearly different in kind, possibly even at odds. If the kind of pattern we're looking for is the context-relative, sometimes-scientific naturalness of the kind that matters to birders, then how could it also be the complex, psychologically revealing but metaphysically obscure kind of naturalness that unites our use of container words?

_____

[92] See Chalmers, Manley, and Wasserman (2009) for a number of excellent contributions to this literature.

[93] Hirsch (2009).

[94] Sider (2009).

The real problem here is a conflation of two distinct projects. The first project is the empirical, linguistic/psychologistic project of discovering the internal constraints on meaning—the kind of constraints that makes it the case that you can be near a house if you're outside it, but not if you're inside it. The second project consists in describing the normative, context-relative constraints on goodness of concept. (Or goodness of parameter setting, or goodness of polysenym choice, etc.) The first project is a theory of reference: how, in fact, do speakers use their words? The second is a theory of deference: under what circumstances should speakers be willing to adjust that usage? The first explores the empirical facts about speakers' semantic competence. The second explores the normative considerations governing the use to which they put that competence.

Lewis and Weatherson are right to observe that scales of naturalness will play an important role. No one could deny that cutting the world at the joints is a pervasive goal in our linguistic practice. However, I differ with them about two features of those scales. First, those scales should not be thought of as meaning determinative, but rather as underwriting our judgements of goodness of concept. Second, the scales themselves are not always scales of *metaphysical* naturalness. Sometimes they are—as in the context of scientific discourse. But often they are not, or not exclusively, as in the context birdwatching, or whaling, or for that matter, writing a loveletter or driving a car or learning a language or any of the activities we engage in that do not concern to the exclusion of other factors, categories that would appear in a finished physical theory.

The view described in MUMPS thus suggests that a great deal of work in philosophy does and indeed should consist in advancing the empirical project of semantics, internalistically conceived. The way that we talk and the categories that we pick out with our ordinary language are topics of intrinsic interest to philosophers. But MUMPS also suggests that this does not exhaust the philosophical work to be done. There remains the project of figuring out which concepts are best for the uses—practical, scientific, and philosophical—to which we put them. In the context of philosophy, those questions can be tremendously difficult to answer. But again, that difficulty is the right result. In other words, the consequences of MUMPS for the problems of philosophy are

very much like the consequences of MUMPS for the problems of marine biology. The easy questions are still easy and the hard questions are still hard.

Bibliography

Anand, P. and A. Nevins. 2004. "Shifty Operators in Changing Contexts," *Proceedings of Semantics and Linguistic Theory* 14.

Barker, C. 2002. "The Dynamics of Vagueness," *Linguistics and Philosophy* 25, 1–36.

Braddon-Mitchel, D. 1995. "The Subsumption of Reference," *British Journal for the Philosophy of Science*. 56(1): 157-178

Burge, T. 1979. "Individualism and the Mental," in French, Uehling, and Wettstein (eds.). *Midwest Studies in Philosophy IV*, Minneapolis: University of Minnesota Press, pp. 73-121.

Capppelen, H. and J. Hawthorne. 2009. *Relativism and Monadic Truth*. New York: Oxford University Press, Inc.

Chalmers, D., D. Manley, and R. Wasserman. (eds.). 2009. *Metametaphysics*. New York: Oxford University Press, Inc

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, N. 2000. *New Horizons in the Study of Language and Mind*. Cambridge, UK: Cambridge University Press.

Chomsky, N. 2007. "Biolinguistic Explorations: Design, Development, Evolution," *International Journal of Philosophical Studies*. 15(1): 1-21.

DeRose, K. 1991. "Epistemic Possibilities," *Philosophical Review* 100: 581–605.

Ducrot, O. 1972. *Dire et ne pas Dire: Principes de Sémantique Linguistique*. Paris: Hermann.

Egan, A. Forthcoming. "De Gustibus Non Disputandum Est (At Least, Not Always)," Forthcoming in Feldman R. and T. Warfield (eds.). *Disagreement.* Oxford: Oxford University Press.

Egan, A. 2007. "Epistemic Modals, Relativism, and Assertion," *Philosophical Studies*, 133: 1–22.

Egan, A., J. Hawthorne, and B. Weatherson. 2005. "Epistemic Modals in Context," In G. Preyer and G. Peter (eds.), *Contextualism In Philosophy: Knowledge, Meaning, and Truth* (pp. 131–170). Oxford: Oxford University Press.

von Fintel, K. "No Head Injury Is Too Trivial To Ignore," *Semantics Etc*. January 21, 2004. Retrieved July 7, 2009. <http://semantics-online.org/2004/01/no-head-injury-is-too-trivial-to-ignore>.

von Fintel, K. and A. Gillies. 2008. "CIA Leaks," *Philosophical Review*, 117: 77-98.

von Fintel, K. and A. Gillies. Forthcoming. "Might Made Right," in Egan and Weatherson (eds.), *Epistemic Modality*. Oxford, UK: Oxford University Press.

Geurts, 1998. "The Mechanisms of Denial," *Language*, 74(2): 274-307.

Grice, H.P. 1975. "Logic and Conversation," in Cole and Morgan (eds.), *Syntax and Semantics, 3: Speech Acts*. New York: Academic Press. Reprinted in Grice (ed.), *Studies in the Way of Words*. Cambridge, MA: Harvard University Press. 1989. pp. 22–40.

Haiman, J. 1979. "Hua: A Papuan Language of New Guinea," in Shopen (ed.), *Languages and Their Status*. Cambridge: Winthrop Publishers, Inc., pp. 35-90.

Hirsch, E. 2009. "Ontology and Alternative Languages," in Chalmers, Manley, Wasserman (eds.), *Metametaphysics*. New York: Oxford University Press, Inc., pp. 231-259.

Horn, L. 1989. *A Natural History of Negation*. Chicago: Chicago University Press.

Horn, L. 2004. "Implicature," in Horn and Ward (eds.), *The Handbook of Pragmatics*. Malden: Blackwell Publishing Ltd., pp. 1-28.

Horwich, P. 2004. "A Use Theory of Meaning," *Philosophy and Phenomenological Research*. 68(2): 351-371.

Inoue, K. 1979. "Japanese: A Story of Language and People," in Shopen (ed.), *Languages and Their Speakers*. Philadelphia: University of Pennsylvania Press, pp. 241-300.

Jakobson, R. 1956. "The Speech Event and the Functions of Language," in Waugh and Monville-Burston (eds.), 1990. *On Language/Roman Jakobson*. Cambridge: Harvard University Press.

Jespersen, O. 1933. *Essentials of English Grammar*. London: Allen and Unwin.

Kennedy, C. 1999. *Projecting the Adjective: The syntax and Semantics of Gradability and Comparison*. New York: Garland. (1997 UCSC Ph.D thesis).

Lasersohn, P. 2005. "Context Dependence, Disagreement, and Predicates of Personal Taste," *Linguistics and Philosophy*, 28: 643–686.

Lewis, D. 1983, "New Work for a Theory of Universals," Australasian Journal of Philosophy, 61: 343–77.

Lewis, D. 1984. "Putnam's Paradox," *Australasian Journal of Philosophy.* 62: 221-236.

Ludlow, P. 2006. "From Sherlock and Buffy to Klingon and Norrathian Platinum Pieces: Pretense, Contextualism, and the Myth of Fiction," *Nous-Supplement*: *Philosophical Issues*. 16: 162-183.

Ludlow, P. 2008. "Cheap Contextualism," *Nous* . Philosophical Issues: Annual Supplement, Sosa and Villanueva (eds.).

Ludlow, P. Unpublished Manuscript. *Philosophical Issues in Generative Linguistics*.

MacFarlane, J. 2007. "Relativism and Disagreement," *Philosophical Studies* 132: 17-31.

Merrill, G.H.. 1980. "The Model-Theoretic Argument Against Realism," *The Philosophy of Science*. 47: 69-81.

Pinker, S. 2004. *The Language Instinct*. New York: Harper Collins.

Potts, C. 2005. *The Logic of Conventional Implicatures*. New York: Oxford University Press, USA.

Potts, C. 2007. "The Expressive Dimension," *Theoretical Linguistics* 33(2): 165-198.

Price, T. 2007. *Speciation in Birds*. Greenwood Village: Roberts & Co. Publishers.

Pullum, Geoffrey, "Nearly All Strings of Words Are Ungrammatical," *Language Log*. January 20, 2005. Retrieved July 7, 2009. <http://itre.cis.upenn.edu/~myl/languagelog/archives/001843.html>.

Pullum, Geoffrey, "'Everything Is Correct' Versus 'Nothing Is Relevant'," *Language Log*. January 26, 2005. Retrieved July 7, 2009. <http://itre.cis.upenn.edu/~myl/languagelog/archives/001843.html>.

Putnam, H. 1975. "The Meaning of Meaning," *Philosophical Papers, Vol. II : Mind, Language, and Reality*. Cambridge: Cambridge University Press.

Rabin, N. 2009. "Dominant-paradigm-subverting Hippified Case File #137: Even Cowgirls Get The Blues," *The Onion*. May 13, 2009. Retrieved July 10, 2009. <http://www.avclub.com/articles/dominantparadigmsubverting-hippified-case-file-137,27976/>

Roberts, C. 1996 "Information Structure: Towards an integrated formal theory of pragmatics," in Yoon and Kathol (eds.) *OSUWPL Volume 49: Papers in Semantics*. The Ohio State University Department of Linguistics. Available in updated form on the Semantics Archive.

Reimer, M. 1998. "Donnellan's Distinction/Kripke's Test," *Analysis* 58(2): 89-100.

"Road to Rhode Island," *Family Guy*. MacFarlane, S. Fox. May 30, 2000.

van der Sandt, R., and E. Meier. 2003. "Denials in Discourse," Michigan Linguistics and Philosophy Workshop. (ms., University of Nijmegen).

Schlenker, P. 2003. "A Plea for Monsters," *Linguistics & Philosophy,* 26(1): 29-120.

Sider, T. 2009. "Ontological Realism," in Chalmers, Manley, Wasserman (eds.), *Metametaphysics*. New York: Oxford University Press, Inc., pp. 384-423.

Smith, M. 1994. *The Moral Problem*. Oxford: Basil Blackwell.

Stalnaker, R. 1970, "Pragmatics." *Synthese 22*. Also in Stalnaker 1999, ch. 1.

Stalnaker, R. 1978. "Assertion." *Syntax and Semantics*. 9: 315-332. Also in Stalnaker 1999, ch. 4.

Stalnaker, R. 1999. *Context and Content*. Oxford: Oxford University press.

Stalnaker, R. 2001. "On Considering a Possible World as Actual," *Aristotelian Society: Supplementary Volume.* 75 (suppl.): pp. 141-156.

Stalnaker, R. 2008. "The New Relativism," Address to the NYU/Columbia Philosophy Graduate Student Conference.

Stanley, J. 2002. "Modality and What is Said," in Tomberlin (ed.), *Philosophical Perspectives, 16, Language and Mind*. Malden: Blackwell Press. pp. 321-344.

Stephenson, T. 2007. "Judge Dependence, Epistemic Modals, and Predicates of Personal Taste." *Linguistics and Philosophy* 30. 487-525.

Tyson, N. 2009. *The Pluto Files*. New York: W.W. Norton & Co.

Wason, P. C., and Reich, S. S., "A Verbal Illusion," Quarterly Journal of Experimental Psychology 31 (1979): 591-97.

Weatherson, Brian. 2003. "What Good Are Counterexamples?" *Philosophical Studies* 115: 1-31.