

Introduction, transmission, and the foundations of meaning

Jeff Speaks

January 8, 2009

1	The analysis of meaning	1
1.1	The Gricean program	2
1.2	Meaning and belief	7
2	Supervenience and the priority of thought	10
3	World history supervenience and world slice supervenience	13

In his 1956 correspondence with Roderick Chisholm, Wilfrid Sellars held the view that “the metalinguistic vocabulary in which we talk about linguistic episodes can be analyzed in terms which do not presuppose the framework of mental acts”, in part because “the categories of intentionality are nothing more nor less than the metalinguistic categories in terms of which we talk epistemically about overt speech as they appear in the framework of thoughts.”¹

The last half century has not been kind to this view. Instead, the vast majority of philosophers of mind and language have sided with Chisholm, who held that we can “explicate the intentional characteristics of language by reference to believing and to other psychological attitudes.”² By far the most widely held and well-worked out views of the foundations of facts about the meanings of expressions in public languages take these facts to be reducible to, or constituted by, or explained by, facts about the propositional attitudes — for example, the beliefs, desires, and intentions — of speakers of those languages. Proponents of this sort of view hold to the thesis that facts about mental content are prior to facts about public language meaning; for convenience, I’ll call these views *mentalist* theories of meaning.

An initial question about mentalism asks what this thesis means — can we explain what ‘reducible to’ or ‘constituted by’ or ‘explained by’ means in this sort of case?

1 The analysis of meaning

One reasonably clear answer to this question is that the mentalist claims to provide an *analysis* of the facts about meaning in terms of facts about mental content — an analysis of what it is for a fact of the former sort to obtain in terms of facts of the latter sort. To

¹See Chisholm and Sellars (1958).

²Of course, not everyone has agreed with Chisholm; important dissenters include Gauker (1994), McDowell (1994), and Brandom (1994).

give an analysis of this sort is to provide metaphysically necessary and sufficient conditions for an expression to mean such-and-such in terms of some collection of facts about mental content, and to defend the view that these conditions are, in some sense or other, more fundamental or basic than the facts about meaning to be analyzed. I'll call this view *analytic mentalism*.

So put, the analytic project presupposes rather than provides an understanding of 'priority'; but this needn't be too serious a problem. We might have some grip on what explanatory priority means, and be able to recognize cases in which one sort of fact is prior to another, without being able to give an analysis of the relation of explanatory priority itself.

Most mentalists have engaged in the analytic project, in this sense. The first question which any analytic mentalist must answer is: *Which* facts about mental content get into the analysis? The two most well-developed versions of the analysis of meaning in terms of mental content differ in their answers to this question. The first, Gricean analysis of meaning focuses on the communicative intentions of speakers of the language, while the second analysis of meaning focuses on the beliefs of speakers of the language. Though consideration of every variant of these two sorts of analyses of meaning is beyond the scope of this paper, in what follows I will argue that each of these versions of analytic mentalism face obstacles whose significance has not been fully appreciated.

1.1 The Gricean program

Grice thought that (1) facts about what expressions mean are to be explained, or analyzed, in terms of facts about what speakers mean by utterances of them; and he thought, further, that (2) facts about what speakers mean by their utterances can be explained in terms of the intentions of speakers. These two theses comprise the 'Gricean program' for reducing meaning to the contents of the intentions of speakers. In what follows I'll focus on part (2) of the Gricean program, the attempt to explain speaker-meaning in terms of audience-directed intentions.

The basic version of the Gricean analysis of speaker-meaning in terms of the intentions of speakers is the following:

- [G] a means p by uttering $x \equiv a$ intends in uttering x that
- (1) his audience come to believe p ,
 - (2) his audience recognize this intention, &
 - (3) (1) occur on the basis of (2)³

One way to see the intuitive motivation behind Grice's analysis is to begin with the idea that meaning something by an utterance is a matter of trying to convey one's beliefs. Trying to convey one's beliefs can be thought of as intending someone to share one's beliefs; but, fairly clearly, you can intend by an action that someone form a belief that p without meaning p by your action. An example here might help. Suppose I turn to you and say, "You're standing on my foot." I intend that you hear the words I am saying; so I intend that you believe that I have said, "You're standing on my foot." But I do not mean by my utterance that I have said, "You're standing on my foot." That *is* my

³See Grice (1957, 1969). For reasons we will not go into, this should be understood as requiring a single, self-referential intention, rather than three separate intentions. [cite harman] Obviously, [G] won't, without modification, handle non-assertoric sentences like imperatives and questions; in what follows I ignore this sort of problem.

utterance — what I mean by it is the proposition that you are standing on my foot, or that you should get off of my foot. I do not mean by my utterance that I am uttering a certain sentence.

This sort of example indicates that speaker meaning can't just be a matter of intending to cause a certain belief — it must be intending to cause a certain belief in a certain way. But what, in addition to intending to cause the belief, is required for meaning that p ? Grice's idea was that one must not only intend to cause the audience to form a belief, but also intend that they do so on the basis of their recognition of the speaker's intention. This condition is not met in the above example: I don't expect you to believe that I have uttered a certain sentence on the basis of your recognition of my intention that you do so; after all, you'd believe this whether or not I wanted you to. This is all to the good.

However, even if [G] can be given a fairly plausible motivation, and fits many cases rather well, ultimately it seems unlikely that there is any connection between speaker-meaning and intended effects stable enough to ground an analysis of the sort that Grice envisaged. To see why, focus first on clause (1) of [G], the requirement that the speaker intend that his audience come to believe the proposition that he means by his utterance.

There are many cases in which we speak meaningfully without intending that our audience form the relevant belief, as when we remind someone of something — in such cases, the audience already has the belief, and so can't come to have it. An initially plausible fix is to change (1) from 'believes' to 'occurently believes' or 'actively believes.' This may help with cases of reminding, but has no obvious application to other counterexamples to (1). As Grice himself noted, instances of confessing something to someone who's caught you red-handed, or taking an oral exam, are both cases in which we mean something by our utterances without intending that our audience come to occurently believe the proposition we mean.⁴ In both cases, the audience already has the relevant occurent belief.

Perhaps the Gricean should modify (1), not by focusing on the distinction between occurent and non-occurent beliefs, but rather by changing his view of the content of the belief which the speaker must intend to cause. One idea along these lines is that when I mean p by an utterance, I don't always intend that my audience come to believe p , but do always intend that my audience come to believe that *I believe that p* .⁵ Clause (1) of the account could be modified accordingly. But this doesn't address obvious variants of the above cases. Suppose, for example, that my daughter confesses to having given her pet fish too much food. In this case, she knows that I already believe that she gave the fish too much food, *and* that I already believe that she believes this. (She might also know that I occurently believe these things.) Nonetheless, she might still mean by her utterance that she gave the fish too much food.

The idea that meaning p is partly a matter of intending that an audience believe that the speaker believes p also leads to some other unattractive consequences. It seems clear that in some cases, I can mean by an utterance that I believe something, without also meaning the proposition which is the content of that belief. For example, if I say "I believe that Indiana is the cultural capital of the United States" I might mean to say something just about my own beliefs; I might have no intention at all of meaning by my utterance that Indiana is in fact the cultural capital of the United States. In this case I will likely intend that my audience come to believe that I believe that Indiana is the cultural capital of the United States. But then I will, on the present view, satisfy the

⁴See among other places Grice (1969).

⁵See Grice (1969), 111.

conditions for meaning by my utterance that Indiana is the cultural capital of the United States (since I intend that my audience come to believe that I believe this proposition) — and this contradicts our assumption that I meant by my utterance only that I believe this proposition.

The Gricean might at this point turn to more drastic revisions; perhaps the key to speaker meaning is not the intention to bring about a belief, but rather the intention to bring about some other sort of mental state. For example, perhaps, as Neale (1992) has suggested, the relevant intentions are intentions to cause agents to *entertain* certain propositions.

There are two problems here. The first is that the modification leads to yet more counterexamples. Consider, for example, utterances of disjunctions. Speaker-meaning clearly does not distribute over disjunction — I can mean by my utterance that either arithmetic is incomplete or the moon is made of cheese without meaning by my utterance that the moon is made of cheese. However, I might never intend someone to entertain a disjunction without intending them to entertain the disjuncts — I might, after all, realize that it is impossible for someone to do the former without also doing the latter.⁶

The second problem is that this modification to clause (1) makes clause (3) impossible to satisfy in many cases. Usually when you hear a certain sentence which in the context means p , you immediately, almost reflexively, entertain the proposition p ; this is the condition that normal speakers are in with respect to sentences they understand. Suppose, I, the speaker, know this. Then I can hardly intend that you come to entertain p *on the basis of your recognition of my intention* when I utter this sentence, as clause (3) requires — I know that you will entertain this proposition irrespective of your recognition of this intention.

The Gricean might reply (as Neale does) that clause (3) is in any case an implausible condition on speaker-meaning, so that dropping this clause from the account is not much of a cost. To get a sense of the problems caused by (3), consider Grice's own example of an utterance of the conclusion of an argument. Clause (3) of [G] requires for a speaker to mean p not only that she intend that her audience come to believe p and that her audience recognize this intention, but also that her audience come to this belief *on the basis of that recognition*. The problem is that, when giving an argument of some kind, a speaker will typically intend that her audience come to believe the conclusion on the basis of belief in the premises, and *not* on the basis of a recognition of the speaker's intention.⁷

Further, if we pursue this suggestion and drop clause (3) from the Gricean analysis, this effectively reduces speaker meaning to the intention to simultaneously cause two effects

⁶A possible fix here is to appeal to the notion of a 'primary intention', in Schiffer's sense: "To specify one's primary intention in doing X . . . is to give one's reason for doing X ," as opposed to specifying an intention which one has because one has a given primary intention (Schiffer (1972), 62). It might be one's primary intention, in this sense, that one's audience entertain a disjunctive proposition and not that they entertain one of the disjuncts; one might intend the latter only because one intends the former. But this modification of the Gricean theory seems too restrictive: Suppose that your car runs out of gas, and I say to you, "There's a gas station around the corner." Clearly, though I do intend that you come to believe that there is a gas station around the corner, I only have this intention because I intend you to come to believe that you can fill your car up around the corner; so the former is a secondary intention and not, in the relevant sense, the reason for my utterance. Nevertheless, it does seem as though I meant by my utterance that there is a gas station around the corner.

⁷Another class of counterexamples to clause (3) of the account, though a less significant one, comes from a sub-class of the class of cases of reminding, discussed above. Suppose that you cannot remember your friend's name, though you feel as though it is on the tip of your tongue. In such a case, I can't intend that you come to form a belief about the name on the basis of recognition of my intention; I know that hearing the name alone will cause you to form the belief.

in an audience: a mental state, and the belief that the speaker intends that the audience come to be in that mental state. But, as an example discussed above shows, this can't be right. Remember the case of my utterance of "You're standing on my foot." As noted above, given that I want you to hear me, I will intend you to believe that I am uttering this sentence. But given that I also want you to know that I am uttering this sentence intentionally, I'll want you to recognize that I intend that you believe this. But this means that the proposition that I have uttered the sentence "You're standing on my foot" is such that (1) I intend that you believe it, and (2) I intend you to recognize this intention. If this were enough for speaker meaning, I would mean by my utterance that I have uttered the sentence, "You're standing on my foot." But I don't — I mean that you are standing on my foot and that you should get off my foot, but I don't mean anything about my own utterances. So again the Gricean faces a dilemma: given the present problem, condition (3) must be included in [G]; but, as noted above, if (3) *is* included, the possibility of meaning something by statements of the conclusions of arguments looks puzzling.

This dialectic of revision and counterexample could be continued.⁸ And in fact there is a persistent thought to the effect that, even if we have not yet arrived at the right form of the Gricean account, *some* true version is out there to be discovered. The discussion of the problems raised by revising clause (1) of Grice's original account has been designed to discourage this thought: each counterexample to a version of the Gricean analysis leads to a revision, which in turn opens the door to new counterexamples. We don't have here a progressive narrowing of problem cases; rather, each version of the analysis creates as many problems as it solves. For this reason, even the range of cases considered so far seems to indicate that while (of course) in many cases there is a connection between what a speaker means by an utterance and the mental states they intend to cause in their audience, there is no connection of this sort stable enough to ground an analysis of speaker-meaning in terms of intended effects.

Nonetheless, it would be good to have a general argument against Gricean analyses, rather than a piecemeal argument by counterexamples. I think that such an argument can be given, by exploiting a variant on Moore's Paradox. Moore drew attention to the oddness of uttering sentences of the form

[M] *S*, but I do not believe that *S*.

While there are contexts with respect to which instances of [M] are true, it is, as Moore pointed out, difficult to imagine a scenario in which a speaker who understands this sentence could utter it sincerely and seriously. The same can be said of a similar sort of sentence, which replaces 'do not believe' with 'do not intend that you will believe':

[I] *S*, but I do not intend you to believe that *S*.

As with Moore's original paradoxical sentence, it is not easy to imagine cases in which a speaker could naturally use a sentence of the form of [I] seriously and sincerely in conversation. But there clearly are cases in which a speaker could utter a sentence which entails an instance of [I] seriously and sincerely. Consider for example, sentences of the form

[K] *S*, but I know that you will not believe that *S*.

⁸Indeed, I've skipped some of the most notorious problem cases, those to do with audienceless uses of language, and especially uses of language in thought. See for discussion Harman (1974) and Chomsky (1975).

Unlike the first two, it is easy to imagine cases in which instances of [K] are uttered seriously and literally. Consider, for example the following:

An unfaithful husband is found out by his wife; he might deeply regret his act and love his wife, but know that there is no way to convince her of this. He says to her, “I love you, though I know you won’t believe that.”⁹

Not only might the husband say this seriously and literally; what the husband says might well be true: it might well be the case that he really does love her, and also be the case that he knows that she will not believe that he does.

What follows from the fact that the husband is speaking seriously and literally, as well as truly? From the fact that he is speaking seriously and literally it follows that among the things he means by his utterance of this sentence is the proposition expressed by the sentence in the context: the proposition that he loves his wife, and knows that she will not believe that he loves her. Moreover, it seems that speaker-meaning distributes over conjunction;¹⁰ it is impossible for a speaker to mean that p and q without also meaning that p . (Just as it is impossible to assert, or believe, a conjunctive proposition without asserting, or believing, the conjuncts.) From this it follows that the husband not only means by his utterance that he loves his wife and knows that she will not believe that he loves her, but also (the first conjunct) that he loves her. So,

(a) The speaker means that he loves his wife by the utterance.

But recall that in the above case the speaker not only spoke seriously, but also truly. So the conjunctive proposition that he loves his wife, but knows that she will not believe this, is true. So both of the conjuncts are true, and, in particular, the second is true:

(b) He knows that his wife will not believe that he loves his wife.

The problem for the Gricean is that it seems plausible that *knowing* that some event will not occur as the result of one’s action seems inconsistent with intending that the event occur: one cannot intend to bring about effects which one knows will not obtain.¹¹ So, from (b) it follows that

(c) He does not intend that his wife will believe that he loves his wife.

But from (a) and (c) it follows that there is some proposition — namely, the proposition that he loves his wife — which is such that the speaker means this proposition by his utterance, and yet he does not intend that his audience believe this proposition. But the Gricean analysis entails that there can be no proposition with both of these properties; so the Gricean analysis must be false.

Of course, we have by now discussed a number of other cases in which are also counterexamples to [G] and its variants — what makes this one special? The interest of

⁹Thanks to Jonathan Beere for this example.

¹⁰That is, indexicality aside, a sentence of the form ‘ $\lceil\alpha$ means that σ and $\sigma'\lceil$ entails ‘ $\lceil\alpha$ means that $\sigma\lceil$ and ‘ $\lceil\alpha$ means that $\sigma'\lceil$.

¹¹If you don’t think that knowledge places this kind of constraint on intending, the example can be changed to some other epistemic state which might seem more clearly inconsistent with intention. For example, the husband might have easily have said: ‘I love you, but I am completely certain that you will not believe me.’ So it seems that the Gricean can only avoid the argument by claiming that there are no epistemic constraints on intention at all, which seems implausible.

the present argument is that it is more than an isolated counterexample: it amounts to a recipe for generating counterexamples in response to various revisions of the Gricean analysis. For example, suppose that a Gricean changes the propositional attitude specified by clause (1) of the analysis to some other propositional attitude, R . Then all that is needed is an example of a sentence analogous to the form of the one uttered by the husband — S , but I know that you will not believe that S — to one of the form: S , but I know that you will not bear R to the proposition that S . It seems plausible that, for any propositional attitude, we'll be able to find a sentence and context such that the sentence may be seriously and truly uttered in that context. So this is a class of counterexamples with substantial generality; it seems likely that, whichever audience-directed intention is seized upon by a particular version of the Gricean analysis, we could construct variants of [K] appropriate to refute that analysis which can be uttered seriously and literally, as well as truly, by speakers.¹²

1.2 *Meaning and belief*

So the Gricean analysis of meaning in terms of speaker-meaning, and speaker-meaning in terms of intentions, faces serious problems. But there's no reason why the analytic mentalist should have to focus on communicative intentions; the analytic mentalist can reduce meaning to any propositional attitude. And while most attention since the 1960's has focused on the Gricean analysis and its descendants, there is at least one other very natural option for the analytic mentalist to pursue. There seems to be a tight connection between the *meaning* of a sentence and the contents of the *beliefs* that speakers have when they utter the sentence seriously. In particular, it seems to be a necessary truth that if a speaker sincerely utters a sentence in a context in which the sentence expresses a certain proposition, the speaker must believe that proposition. Why not try to use this necessary connection between meaning and belief to analyze facts about linguistic meaning in terms of the contents of the beliefs of speakers?¹³

The burden of such a theory will be to say what it takes for a sentence to inherit its meaning from the content of a particular belief. The simplest attempt at doing this is to say that a sentence is correlated with a belief just in case were an agent to accept (endorse, assert) the sentence, then the agent would form the relevant belief. A first try at formulating this sort of analysis of meaning might be something like the following:

$$[B] \quad (S \text{ means } p \text{ in a community } C) \equiv \forall a(a \text{ is a member of } C \rightarrow (a \text{ accepts } S \square \rightarrow a \text{ believes } p))$$

There are delicate issues about the interpretation of the counterfactual on the right-hand side of [B]. It is tempting to argue that it is falsified by sentences like “0=1” which are so obviously false that everyone who understands them rejects them. Consider the nearest possible world in which I accept the sentence “0=1”. Isn't it a world in which the sentence has a different meaning more similar to the actual world than one in which I believe, absurdly, that zero and one are identical? If so, [B] yields the incorrect result that “0=1” does not express the proposition that zero equals one.

However, it is plausible that counterexamples of this sort can be blocked by stipulating away ‘backtracking’ evaluations of the counterfactual in [B], in which one rejects

¹²I think that this sort of argument can also be used against neo-Gricean analyses, like that defended in Davis (2002), but that is beyond the scope of this paper. I hope to develop this argument further elsewhere.

¹³I think that Ramsey (1927) had a view of this sort in mind. See also Loar (1981).

a counterfactual on the grounds that, had the event mentioned in the antecedent of the counterfactual occurred, something in the past (relative to that event) would also have been different, where this latter makes the consequent of the counterfactual false.¹⁴

The real problem with [B] is that facts about the meanings of sentences are, in a certain sense, more fine-grained than facts about the beliefs of agents; and this makes it difficult to give an account of the former in terms of the latter.

Belief has the feature attributed above to speaker-meaning: it distributes over conjunction.¹⁵ Now consider a conjunctive sentence S whose meaning is the conjunction of the propositions p and q . Then each of the following claims will be true:

$$\forall a ((a \text{ is a member of } C) \rightarrow (a \text{ accepts } S \square \rightarrow a \text{ believes } p \& q))$$

$$\forall a ((a \text{ is a member of } C) \rightarrow (a \text{ accepts } S \square \rightarrow a \text{ believes } p))$$

$$\forall a ((a \text{ is a member of } C) \rightarrow (a \text{ accepts } S \square \rightarrow a \text{ believes } q))$$

The problem is that it follows from [B], along with the truth of these three claims, that S is three ways ambiguous: S means $p \& q$, S means p , and S means q . After all, if [B] is to be true, then anyone who accepts S must believe the conjunctive proposition $p \& q$; but the distribution of belief over conjunction entails that they will also believe p and believe q . So all three beliefs are correlated with the conjunctive sentence; and if meaning is defined in terms of correlated beliefs, the sentence has as its meaning each of these three propositions. But this is incorrect: conjunctive sentences are not, in general, three ways ambiguous just by virtue of being conjunctions.¹⁶

While this does show that [B] is false, it seems that there is an easy way to revise it in answer to this objection. We can limit the account of meaning to non-conjunctive sentences, and give a separate account of the meanings of conjunctive sentences in terms of the meanings of their parts.

But this is only a superficial solution; there are counterexamples similar in form to those based on the distribution of belief over conjunction even for simple sentences. To generate such a counterexample, one needs only a sentence S and two non-conjunctive propositions p , q such that every speaker in the relevant community who would believe p upon accepting S would also believe q . For simplicity, I'm focusing on cases in which believing q is a metaphysically necessary consequence of believing p ; but the class of pairs of propositions which can generate counterexamples to [B] is much larger than this.

But even if we restrict ourselves to cases in which having one belief strictly entails having another, there are plenty of plausible examples. Consider, for example, the proposition that Lassie was a brown dog (or collie, or puppy) and the proposition that Lassie

¹⁴For discussion, see Lewis (1979). In the present case, it is natural to think that, if I had accepted "0=1", this could only have been because "0=1" expressed some non-trivially-false proposition; and this latter consideration is what makes it false that, in the world under consideration, I believe that 0=1. But, by ruling out backtracking evaluations of this sort, perhaps we can put cases like this to the side. A residual worry here is that in making this stipulation, we are illegitimately building facts about what sentences mean into the class of facts held fixed for purposes of evaluation of the counterfactual. I am unsure whether this is a genuine circularity.

¹⁵That is, it is a necessary consequence of the truth of a sentence of the form $\ulcorner \alpha \text{ believes that } \sigma \text{ and } \sigma' \urcorner$ that, in that context, the following sentences are also true: $\ulcorner \alpha \text{ believes that } \sigma \urcorner$ and $\ulcorner \alpha \text{ believes that } \sigma' \urcorner$.

¹⁶Actually, matters are a bit worse than this for [B]. Normal cases of ambiguous sentences are ones in which the sentence is sometimes correlated with one belief, and sometimes with another; 'I am going to the bank' is sometimes correlated with beliefs about financial institutions, and sometimes with beliefs about the sides of rivers. But in this case we have a sentence which is always correlated in the relevant way with three beliefs. So the result is not just that conjunctive sentences are ambiguous; it's that they are ambiguous in a heretofore unimagined way. Thanks to Tom Feeney for pointing this out.

was a dog. Is it possible for an agent who accepts, and therefore understands, the sentence “Lassie was a brown dog,” and who believes that Lassie was a brown dog, not to believe that Lassie was a dog? If not, [B] entails that “Lassie was a brown dog” is ambiguous between meaning that Lassie was a brown dog and meaning that Lassie was a dog; this is clearly incorrect. Colors provide a similar example — is it, e.g., a necessary truth that if someone believes that an object is scarlet, she also believes that the object is red? If so, then we can generate a similar counterexample in the obvious way. Another such class of cases exploits the trivial equivalence — paradoxical cases aside — of a proposition p and the proposition which attributes truth to p . Suppose that an agent accepts the sentence “It is true that Lassie was a dog”, and so by the disquotational principle believes that it is true that Lassie was a dog. Is it possible that such an agent could yet fail to believe that Lassie was a dog? Again, it seems unlikely.

One response to these kinds of cases is to try to come up with some criterion for separating simple from complex sentences, and to hold that (i) [B] is only applicable as a theory of meaning for simple sentences, (ii) the meanings of complex sentences are a function of the meanings of the simple sentences, and (iii) all the above cases turn on illicit applications of [B] to complex sentences. This strategy faces two main challenges. First, we need some criterion which classes together all of the sentences which generate counterexamples of the above form; second, we need this criterion to be loose enough to count as simple enough sentences to provide the resources to supply the missing theory of meaning for complex sentences. The prospects for finding a criterion to meet both constraints do not appear to me to be good.¹⁷

It is tempting to try to solve this problem by appealing to a distinction between beliefs gained ‘directly’ as a result of accepting a sentence, and beliefs gained ‘indirectly.’ In one sense, this is clearly correct; if S means p , then there is, intuitively, a more direct link between the acceptance of S in by an agent and that agent’s believing p than between the agent’s acceptance of S and the agent’s coming to believe any other propositions. The problem, though, is not to find intuitions of directness; the problem is to explain what these intuitions are tracking in a noncircular way. One can’t, after all, appeal to causal differences; it is not as though, in the sorts of cases described above, the agent comes to believe p and that an effect of this is their coming to believe q — in the cases described above, the former condition is metaphysically sufficient for the latter.¹⁸

¹⁷Both challenges seem to me to pose serious problems. I am reluctant to think of ‘is scarlet’, for example, as being syntactically complex; and even if it were, I am unsure how the assignments of meanings to simple sentences could be extended to a complete theory which applies to complex sentences as well.

A related move for the mentalist is to try to solve the problem via a distinction between the various propositions which are candidate meanings for a sentence. One proposal which several people have suggested to me is that, if it is a necessary truth that someone who accepts a sentence S believes a set of propositions $p_1 \dots p_n$, we should assign as the meaning of S the strongest of these propositions. This works nicely for some of the cases discussed above, such as the conjunctive sentence. Because the conjunctive proposition entails but is not entailed by its conjuncts, the present proposal assigns, correctly, the conjunctive proposition as the meaning of the the conjunction. The solution isn’t general though, as is shown by a small variant on the preceding one: if one of the conjuncts is a necessary truth, then neither the conjunction nor the other conjunct will be uniquely strongest; but, for all that, the conjunction is not ambiguous. The example of a proposition and the proposition which ascribes truth to it is similar. And there are some cases in which the weaker of two propositions is the right interpretation. For example, consider the sentence ‘I am actually a teacher.’ I utter this only when I believe that I am actually a teacher — i.e., a teacher in the actual world — but whenever I believe this, I also believe that I am a teacher. Even so, the sentence ‘I am actually a teacher’ does not mean that I am a teacher; it means that I am a teacher in the actual world.

¹⁸Another possibility is to explain the direct/indirect distinction in terms of sentences in an internal language of thought, or in terms of an internalized semantic theory. While this is a revision of [B] worth

It is important here not to be so focused on the counterexamples as to lose sight of the source of the objection. The basic point is that it seems to be in the nature of belief that facts about the beliefs of agents come, so to speak, in clumps: for many propositions, it is a necessary truth that any agent who believes such a proposition must also believe one or more distinct but related propositions. But facts about meaning just are not ‘clumpy’ in this way: there are no propositions which are such that it is a necessary truth that, if a sentence has it as its meaning, it must have one or more distinct propositions as its meaning as well. Because facts about meaning are, in this sense, more fine-grained than facts about belief, it is difficult to see how we could give an account of the former in terms of the latter.¹⁹

2 Supervenience and the priority of thought

Of course, further variations on [G] and [B], our two versions of analytic mentalist theories of meaning, are possible. But at this point I want to shift attention to another sort of route that the mentalist might take. One might well think that the mentalist as such needn’t be overly worried by these difficulties with the analytic project — any more than one who thinks that physical facts are explanatorily prior to mental facts need be worried, as such, about difficulties in giving an account of what it is to feel pain in terms of purely physical properties. Such cases suggest that we can have explanatory priority without analysis; so why not just hold to the priority of mental content over meaning without holding any commitment to the success of the analytic project?

This raises two related questions for the mentalist. First, how can we establish the relevant priority claim without providing an analysis of meaning in terms of facts about mental content? And, second, if the mentalist is not claiming that a given fact about linguistic meaning can be analyzed in terms of facts about mental content, then what does the claim that mental content is prior to meaning amount to? A plausible and popular answer to this pair of questions is that saying that the facts about mental content are prior to the facts about meaning is saying that the facts about meaning are ‘nothing over and above’ the facts about mental content, and that we can argue for this by showing that the facts about meaning supervene (in a sense to be specified) on the facts about mental content. Let’s call this loosely defined position *supervenience mentalism*.

What sort of supervenience of meaning on mental content would be sufficient to show the priority of the latter over the former? Let’s define the supervenience of one class of properties on another as follows:

The A properties supervene on the B properties \equiv_{af} for any two worlds w and w^* , every one-to-one mapping of individuals in w and w^* such that each individual in one world alike with respect to the B -properties of the individual

pursuing, it’s beyond the scope of this paper. A short reply to this view is that it seems implausible to say that, if we discovered that we were all possessed of internalized semantic theories which mapped conjunctions in English onto each of their conjuncts as well as a conjunctive sentence of the language of thought, we would then conclude that conjunctions were three-ways ambiguous.

¹⁹This sort of worry is also a problem for a familiar sort of variant on [B], which attempt to analyze meaning in terms of correlated beliefs along with mutual beliefs by speakers about those beliefs. Lewis (1975), among others, gives a theory of this sort. But we can imagine a linguistically sophisticated community, in which, for example, all speakers know — and know that each other know — that belief distributes over conjunction. (Or fill in your favorite of the above examples.) For such a community, the examples above will work as well against a convention-theoretic account of Lewis’ sort as they do against the simpler [B].

with which it is paired in the other world is also such that the paired individuals are alike with respect to their *A*-properties.²⁰

Clearly, we can't say that supervenience in this sense is sufficient for priority; priority is an asymmetric relation, but supervenience in this sense is not. But we can use the above definition of supervenience to formulate an asymmetric relation which might appear to be a good indicator of priority:

Asymmetric Supervenience

The *B* properties are prior to the *A* properties iff (1) the *A* properties supervene on the *B* properties, and (2) the *B* properties don't supervene on the *A* properties.

Intuitively, the idea is that once we fix the *B*-properties in a world, we've thereby determined the distribution of *A*-properties, but that fixing the distribution of the *A*-properties still leaves open some questions about which individuals instantiate which *B*-properties. This seems to afford the *B*-properties a kind of privileged position with respect to the *A*-properties; it seems to show that the *B*-properties are at least in part independent of the *A*-properties, whereas the *A*-properties are determined by the *B*-properties.

Moreover, it seems that if the Asymmetric Supervenience Test is a good one, things look pretty good for supervenience mentalism. Remember, the supervenience mentalist claims that facts about the beliefs, intentions, desires, and other propositional attitudes of subjects are prior to facts about what the words of their language mean. So, let's focus on the following two sets of properties:

MC: properties of standing in certain propositional attitude relations, like believing, desiring, or intending, to contents.

S: properties related to the semantics of one's language, like speaking a language in which such-and-such expression has such-and-such a meaning.

It certainly seems as though *S* asymmetrically supervenes on *MC*. It is quite difficult to see how two worlds could be alike with respect to all of the intentions, beliefs, and other mental states of every individual, but different with respect to the meaning of some linguistic expression; but it's easy to imagine a pair of worlds mentally alike but for the fact that in one Bob has a passing desire for a cupcake which are nonetheless identical with respect to the semantic properties of the expressions of every language spoken in those worlds.

But the Asymmetric Supervenience Test can't, as it stands, be sufficient for priority. In most cases in which we're interested, the relevant supervenience relations will not hold between a pair of properties, but between a pair of *classes* of properties. For example, in the present case we're interested in the supervenience relations between mental properties and properties to do with the meanings of words in the languages of speakers. But if we take the Asymmetric Supervenience Test as our guide, whether one property is prior to another seems to depend on apparently arbitrary decisions about how we group properties into classes.

²⁰This is strong global supervenience, in the sense of Bennett and McLaughlin (Fall 2006) and Stalnaker (1996). I'm focusing on global supervenience to make room for the view, shared by many mentalists, that the meaning of expressions in one's lexicon depends not just on the contents of one's own mental states, but also those in one's linguistic community. For present purposes, I'm not sure that much hangs on the distinctions between different sorts of global supervenience claims.

Suppose, for example, that the *A* properties asymmetrically supervene on the *B* properties. Now consider some class *C* of properties on which the *B* properties asymmetrically supervene. If the *B* properties asymmetrically supervene on the *C* properties, they'll also asymmetrically supervene on the union of the *C* properties and the *A* properties. So are the *B* properties prior to the *A* properties, or not? If all we have to go on is the Asymmetric Supervenience Test, then it seems that the *B* properties are prior to the *A* properties, but less fundamental than the *A* properties together with the *C* properties. But this can't be right; the facts about what is ontologically prior to what can't depend on our decisions about which properties to consider together.

It is not easy to know how to reform the Asymmetric Supervenience test to block this sort of result. An initial thought is simply to add the following condition to our test:

(3) There is no class of properties *C* such that the *B*-properties asymmetrically supervene on the union of the *A*- and *C*-properties.

But this would be much too strong; as long as the *B*-properties asymmetrically supervenes on some set of properties or other, there will be such a class *C*. A more refined test would be to append, not (3), but

(3*) There is no class of properties *C* such that the *B*-properties asymmetrically supervene on the union of *A*- and *C*-properties but do not asymmetrically supervene on the *C*-properties alone.

The problem for the supervenience mentalist is that, if we add (3*) to the Asymmetric Dependence test, it is not clear that the modified test yields the result that *MC* is prior to *S*. After all, her opponent is likely to say that, even if *S* does asymmetrically supervene on *MC*, there is a class of properties such that *MC* asymmetrically supervenes on the union of that set with *S*, even though *MC* does not supervene on that set alone. Consider, for example, two further sets of properties:

D: dispositional properties, like properties of agents being disposed to perform certain actions.

P: properties to do with perceptual representation of one's environment.

One sort of opponent of mentalism is likely to think that *MC* asymmetrically supervenes on the union of *S*, *D*, and *P*; and this is not, on the face of it, an implausible claim. It is hard to imagine two worlds alike with respect to the distribution of semantic, perceptual, and dispositional properties but which differ with respect to the beliefs, desires, or intentions of individuals in those worlds. If this is right, then we have two supervenience claims:

S asymmetrically supervenes on *MC*

MC asymmetrically supervenes on *S&D&P*

There's no inconsistency in affirming each of these two claims about asymmetric supervenience; but there is an inconsistency in taking each of the two to be priority-revealing. And of course in deciding which, if either, is priority revealing, the Asymmetric Supervenience Test is no help. So, from the supervenience facts alone, it is less than obvious that

we get a determinate result as to whether properties in MC are prior to the properties in S .²¹

3 World history supervenience and world slice supervenience

So far, our conclusions have been negative. The project of analytic mentalism faces serious problems, and the supervenience relations between meaning and mental content considered in the previous section don't seem to give us an answer to the 'priority' question. I think that we can take some positive steps toward understanding the relationship between meaning and mental content by considering a certain sort of failure of supervenience; or so I'll argue in this section.

Above we defined the supervenience of A -properties on B -properties as follows:

For any two worlds w and w^* , every one-to-one mapping of individuals in w and w^* such that each individual in one world is alike with respect to the A -properties of the individual with which it is paired in the other world is also such that the paired individuals are alike with respect to their B -properties.

Since the properties we're interested in are ones that individuals have at some times but lack at others, one might wonder how the role of time in this formulation should be understood. Let's say that an *order-preserving* mapping of times in w and w^* is a one-to-one mapping of times such that for any two times x, y in w , x is earlier than y in w iff the time in w^* to which x is mapped is earlier in w^* than the time to which y is mapped. Given this, it is natural to understand the one-to-one mapping mentioned in the above definition of supervenience as constrained by an order-preserving mapping of times: when we say that a certain one-to-one mapping of individuals is such that each pair is alike with respect to their A -properties, what we really mean is that a certain one-to-one mapping of individuals is such that each pair is alike with respect to their A -properties *at each time*, where for them to be alike at each time is for an individual in w to have the same A -properties at t as the individual with whom he is paired in w at the time paired with t by some order-preserving mapping of times.²² Let's call this *world history supervenience*, since to say that one property supervenes on another in this sense is to say that any two worlds alike with respect to the distribution of the subvening class of properties over the course of the history of those worlds are also alike with respect to the distribution of the supervening class of properties over the course of the history of those worlds.

But just as we can define a supervenience relation based on world histories, we can also define a supervenience relation based on world slices: the states of possible worlds

²¹For further objections to the idea that metaphysical dependence or priority should be understood in terms of supervenience, see Bennett and McLaughlin (Fall 2006), §3.5 and Schaffer (2009).

The supervenience mentalist might object to the foregoing by claiming that the properties in P belong with the properties in MC . But of course which properties belong with which is part of what is at issue. The supervenience mentalist might also object to the inclusion of D in the anti-mentalists' supervenience base, on the grounds that intentional actions are ultimately to be analyzed partly in terms of beliefs, desires, and intentions. Plausibly, the anti-mentalists who take facts about beliefs, for example, to be explained in terms of linguistic meaning rather than the other way around is already committed to rejecting this sort of causal-psychological theory of intentional action. But it must be conceded that if some such theory of intentional action turns out to be right, the sort of 'parity' argument given above is harder to make. However, perhaps D could be replaced by the class of dispositions to perform certain bodily movements, rather than intentional actions, in certain circumstances.

²²This is the way that distributions of properties over time are handled in the appendix to Paull and Sider (1992).

at particular times. Thinking of a world slice as a pair of a world and a time, we might define one such supervenience relation as follows:

World slice supervenience

For any two world/time pairs w, t and w^*, t^* , every one-to-one mapping of individuals in w at t and w^* at t^* such that each individual in one world is alike with respect to the A -properties of the individual with which it is paired in the other world is also such that the paired individuals are alike with respect to their B -properties.

It is possible for one property to supervene on another in the world-history sense without supervening in the world-slice sense. In fact, our two sets of properties above — MC and S — are so related. I'll argue this point first, and then turn to a discussion of its significance for our understanding of the relationship between mental and linguistic representation.

Consider the following example:

In 2695, a historian of American sports is going through some old magazines in the archives, and comes across a story about someone named 'Pete Rose.' The story talks about how he was banned from baseball for betting on games, etc. etc.

The historian then leaves the archives and says to a colleague, 'I read about this guy, Pete Rose. He was an interesting character ...' When the historian utters this sentence, it seems that the name 'Pete Rose' has a reference for him, and that it refers to the same man to whom it refers for me and other English speakers familiar with the name: a certain late twentieth-century professional baseball player. And this can be so even if in 2695 there are only a few historians of American sports, and so that, immediately prior to this historian delving into the archives, no one had any thoughts, beliefs, or other propositional attitudes about Pete Rose. However we think of the semantics of names, it seems plausible that if no one has any thoughts about an object, then no one has any thought whose content includes the concept expressed by that name. So, if the name had a reference, and hence a meaning, before being encountered by the historian, it seems that the meaning of the name was not a concept which figured in the thoughts of any person alive at that time.

But does the name have a reference before being encountered by the historian? It is hard to see how it could fail to have a reference. After all, if it did not have a reference in, say, 2694, then the historian would be effectively introducing a new name into the language. And it would then be puzzling how the historian could do this: how could the historian be in a position to introduce a name which genuinely refers to Pete Rose — as the historian's uses of the name seem to? (We can imagine that the archives do not provide uniquely identifying information, if that matters.) To put the same point another way: if the name lacked a meaning when the last person with beliefs about Pete Rose died, and it acquired a meaning in 2695, what gave it the meaning it acquired in 2695? It's hard to see how it could 'automatically' reacquire the same meaning it had unless it simply had that meaning all along.

But if this is right, then meaning does not supervene in the world-slice sense on mental content. For imagine the world described above in 2694, before the discovery. We can imagine a world-slice exactly the same as this one with respect to the mental

representations of each individual in the world, but in which ‘Pete Rose’ either is not a meaningful name, or is the name of someone else. So there can be a difference in the semantic properties of expressions at a time without any corresponding difference in the facts about mental content at that time.

This sort of phenomenon can clearly arise with respect to any sort of name, whether of a person, or a place, event, or something else; it also seems to apply to names for properties, and other abstract singular terms. It is also easily extended to the case of natural kind terms; imagine a case like the above in which the historians of American sports are replaced by historians of the fauna of North America, and ‘Pete Rose’ is replaced by a term for an extinct species, like ‘passenger pigeon.’ The historians could come to have *de re* thoughts about the passenger pigeon, just as the historians in the example above could come to have *de re* thoughts about Pete Rose.

In fact, it seems possible to run this sort example for at least many ordinary predicates. Consider color words. Suppose that over time humans lose the ability to distinguish shades of red and green, so that eventually all language users are red/green color blind. The words ‘red’ and ‘green’ might then, let’s imagine, fall into disuse and be forgotten; and eventually it might become the case that no one had any thoughts involving the concepts expressed by these terms. But the terms might then be rediscovered in a way analogous to the rediscovery of ‘Pete Rose’ described above; those who rediscovered the terms might note that they were used as color terms in a way similar to words familiar to them, like ‘yellow’ and ‘blue’, and eventually learn about the events which brought about the disappearance of ‘red’ and ‘green’ from the language. Couldn’t these people then use sentences like ‘I wonder which things are red, and which things are green’ to express thoughts which are about the colors, red and green? If so, then the argument runs as above: because the relevant language users are in no position to introduce a term to refer to the properties red and green, ‘red’ and ‘green’ must have referred to these properties prior to their rediscovery. But if they referred to these properties, then they must have had a meaning which determined these properties as reference. But, since by hypothesis no one has thoughts about these properties at the time t immediately prior to the rediscovery of ‘red’ and ‘green’, no one has thoughts involving the concepts expressed by these terms at t . And so, as above, it is easy to imagine a world-time pair in which we hold fixed the thoughts of everyone in the world described above at t , but change the reference (and hence meaning) of ‘red’ and ‘green.’

To be sure, this sort of case not as straightforward as the original cases of names and kind terms. But it is anyway a bit awkward to hold that the meanings of ordinary predicates do supervene in the world-slice sense on the facts about mental content, whereas the meanings of names do not. After all, if we think of the intension of a predicate as fixed by a (possibly complex) property, then it is plausible that, for any predicate F , we can introduce an abstract singular term n as a name for the property which determines the intension of F . Then we can use this term to form predicates of the form ‘instantiates n ’ which will have the same intension as F . Sophisticated language users apt to form predicates of this sort will of course regard them, correctly, as sharing their intension with the corresponding predicate. But notice that in the above examples of world slice supervenience failure, it’s not just that no one has any propositional attitudes involving the relevant concept; it’s also true that no one has any propositional attitudes involving any concept which shares its intension with the concept in question. But if this is right, then the failure of world slice supervenience for abstract singular terms entails the failure of world slice supervenience for predicates, given that those predicates share their intensions

with predicates formed using abstract singular terms which we already know fail to world-slice supervene on the facts about mental content.

So, for at least many sorts of linguistic expressions, their meaning at a time does not supervene on the propositional attitudes of users of the relevant language at that time.²³ What does this supervenience failure show us about the relationship between meaning and mental content?

For one thing, this supervenience failure amounts to what Brian McLaughlin has called a refutation by appeal to a FIST (false implied supervenience claim) of standard versions of analytic mentalism.²⁴ Standard analyses of meaning of either the Gricean or belief-based sort analyze the meaning of a word at t in terms of the propositional attitudes of users of the language at t ; in this way, at least, such theories have to be revised.

Second, and perhaps more importantly, the failure of world-slice supervenience puts some pressure on the supervenience mentalist's claim that, in some sense or other, the facts about meaning are 'nothing over and above' the facts about mental content. After all, in all of the standard examples of cases in which the supervenience of one sort of property on another is taken to show that instantiation of the supervening property is 'nothing over and above' instantiation of the subvening property, we have no gap between world-history supervenience and world-slice supervenience. Consider the following examples:

The supervenience of the shapes presented on a TV screen on the colors of the pixels.

The supervenience of shape properties on microphysical properties.

The supervenience of phenomenal character on intrinsic physical properties.

The supervenience of dispositional properties on categorical properties.

One might of course dispute some of these supervenience claims; but what seems clear is that *if* these supervenience relations hold in the world-history sense, they also hold in the world-slice sense.

But these negative points leave unanswered the more interesting question: what does the fact that meaning supervenes on mental content over world histories but not world slices show us about the relationship between meaning and mental content? We can get some guidance from other instances of world-history supervenience without world-slice supervenience:

The supervenience of the laws in effect in a democracy on the actions of legislators.

The supervenience of the facts about artworks on the intentions and actions of artists.

Plausibly, two world-slices could be alike with respect to the actions of legislators in those worlds, but different with respect to which laws are in effect; but this will always be due to some difference in legislator actions at some previous time in the history of the world from which the slice is taken. This indicates that the first of these two supervenience claims holds in the world history sense, despite failing in the world slice sense. Parallel remarks apply to the supervenience of art works on the intentions and actions of artists.

²³We needn't assume that this holds for *all* expressions; it is difficult to construct an analogous case for logical constants.

²⁴See McLaughlin (1984).

In each case, the explanation of why world-history supervenience holds but world-slice supervenience does not is fairly obvious. If we ask what it is for a law in a certain sort of democracy to be in effect, our answer has two parts: we must say what it is for a law to be instituted, or put into effect, and what it is for a law to continue to stay in effect over time. And our theories of law-institution and law-transmission will not, of course, be the same; certain actions of legislators are required to institute a law, but legislators need not be constantly engaging in these actions for the law to stay in force.

A face-value interpretation of the facts about supervenience of meaning on mental content suggests that our answer to the question, ‘What is it for a word to have a given meaning at a time?’ should also have two non-interchangeable parts: an explanation of what it is for a word to be introduced with a given meaning, and of what it is for it to maintain that meaning over time.

This is hardly a new idea. One of the guiding ideas of *Naming and Necessity*, after all, was that in order to understand the reference of names and kind terms at a time, we need to divide the story about how they acquire a reference into an account of how they are introduced with a given reference, and how they maintain that reference over time and across speakers. And, if we accept Kripke’s views about reference, it seems inevitable that we extend them to the case of meaning. After all, given that meaning determines reference, if some other theory of meaning were available, Kripke’s theory of reference would be dispensable. Given the widespread acceptance of Kripke’s ideas, it is surprising that theories of the foundations of meaning have in general *not* taken this two-part structure.

Perhaps the reason for this divergence is that people have thought that the distinction between introduction and transmission was relevant only to the sorts of names and kind terms that Kripke considers. But, first, the examples above indicate that this is a mistake, and that the distinction between introduction and transmission is relevant not just to names and kind terms, but also to standard predicates. And second, names and kind terms are hardly a negligible proportion of natural language expressions; one would think that if theories of meaning were meant to exclude these expressions from their scope, this should be made explicit.

But let’s suppose that the right account of the foundations of meaning will consist of a pair of theories: an account of what it is to introduce a term with a given meaning, and an account of how meaning is preserved, or transmitted, between speakers and over time. How does this bear on the question of mentalism?

An immediate consequence of the sorts of world-slice supervenience failures above is that meaning transmission needn’t involve any speaker being in any propositional attitude state whose content includes the content of the expression in question. This rules out the most obvious sorts of mentalist theories of meaning transmission, and makes room for meaning to have a certain priority over mental content in particular cases: as in the example of ‘Pete Rose’, speakers can be put in a position to have thoughts with certain contents in virtue of their coming into contact with a linguistic expression with that content — even if there is no one who exists at that time who has a mental representation of any sort with that content.

Mentalist theories of term introduction still remain a live possibility; but the focus on term introduction, rather than on simply using a term with a certain meaning, changes *which* sorts of mental representations might be relevant to the determination of meaning. In particular, it makes it more plausible that perceptual states could play a role here. It is very implausible that to *use* a name with a certain meaning, one must have perceptually

represented the referent of the name at some point — we often use names for things we've never encountered in perception. But it is significantly more plausible — though still, of course, controversial — that one can introduce a name for an object only on the basis of perceptual representation of it. Many predicates might also be introduced via perceptual representations of the properties which are their contents.²⁵

Not all predicates can have their meaning explained in this way, since not all properties which are the contents of predicates are the sorts of properties that can be perceptually represented. Perhaps in these cases, the right account of term introduction can be given by a use theory of the sort which has been developed by Paul Horwich. An example of a predicate whose content plausibly cannot be given by a perceptual experience is 'is true.' As Horwich suggests, we might think of the rule for introducing a term with the meaning of our truth predicate as the adoption of a disposition to provisionally accept every instance of the corresponding T-schema.²⁶

This sort of mixed account of term introduction is, of course, just a sketch of a theory. But suppose that some such theory is correct. Would we then say that a mentalist account of meaning has been vindicated, or not?

In one sense, this sort of view clearly captures some core mentalist commitments: there's a clear sense in which it locates the foundations of intentionality in a kind of mental representation — namely, perceptual representation. In this sense, the mentalist view that all linguistic representation is 'derived intentionality' would turn out to be correct.

But, on the other hand, the above discussion suggests that the relationship between mental and linguistic representation might well prove more complicated than mentalist accounts suggest — and not just because of the need to take account of the distinction between term introduction and term transmission. Perhaps we should understand perceptual representation as prior to and explanatory of certain kinds of linguistic representation; but this needn't entail a mentalist view of the relationship between linguistic meaning and the contents of judgements and beliefs. A view of term introduction and transmission of the sort sketched above opens the door to a view of these sorts of cognitive mental states as inheriting their contents from perception on the one hand, and the semantic properties of public language expressions, on the other. If this turns out to be right, then even if something of the mentalist's thesis of the priority of the mental is correct, the mentalist's view of the relationship between meaning and thought is not.²⁷

²⁵This way of putting things assumes a Russellian view of semantic contents; but the Fregean might think of perceptual experiences as providing modes of presentations of objects and properties which could then go on to serve as the contents of names and predicates, respectively. For a plausible argument that in at least some cases it is impossible to introduce a term with a given content without having a perceptual experience involving that content — since otherwise certain sorts of *de re* claims would, implausibly, be knowable a priori — see Soames (2005).

²⁶See, among other places, ch. 2 of Horwich (2005). Horwich, of course, intends his use theory to be a theory of meaning *simpliciter*, rather than a theory of what it is to introduce a term with a certain meaning. But Horwich's theory seems open to the same sort of refutation by a FIST discussed above: the meaning facts at *t* do not globally supervene on the Horwichian use facts at *t*, and so can't be analyzed in terms of them. Each of the examples developed above work as well against Horwich's theory as against the supervenience mentalist.

²⁷Thanks to the participants in my seminar on the metaphysics of meaning in the spring of 2008 at Notre Dame and, for helpful discussion of supervenience, to Fritz Warfield.

References

- Karen Bennett and Brian McLaughlin, Fall 2006. Supervenience. In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- Robert Brandom, 1994. *Making It Explicit*. Cambridge, MA: Harvard University Press.
- Roderick Chisholm and Wilfrid Sellars, 1958. Correspondence on Intentionality. In *Intentionality, Mind, & Language*, edited by Ausonio Marres, 214–248. Urbana, Illinois: University of Illinois Press.
- Noam Chomsky, 1975. *Reflections on Language*. London: Temple-Smith.
- Wayne Davis, 2002. *Meaning, Expression, and Thought*. New York: Cambridge University Press.
- Christopher Gauker, 1994. *Thinking Out Loud: An Essay on the Relation between Thought and Language*. Princeton, NJ: Princeton University Press.
- Paul Grice, 1957. Meaning. In Grice (1989), 213–223.
- Paul Grice, 1969. Utterer's Meaning and Intentions. In Grice (1989), 86–116.
- Paul Grice, 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard U.P.
- Gilbert Harman, 1974. Stephen Schiffer: *Meaning*. *Journal of Philosophy* 71:7:224–229.
- Paul Horwich, 2005. *Reflections on Meaning*. Oxford: Clarendon Press.
- David Lewis, 1975. Languages and Language. 3–35. Reprinted in Lewis (1983), 163–188.
- David Lewis, 1979. Counterfactual Dependence and Time's Arrow. *Noûs* 13:455–476.
- David Lewis, 1983. *Philosophical Papers*, volume 1. Oxford: Oxford U.P.
- Brian Loar, 1981. *Mind and Meaning*. Cambridge: Cambridge University Press.
- John McDowell, 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
- Brian McLaughlin, 1984. Perception, Causation, and Supervenience. *Midwest Studies in Philosophy* 9:569–592.
- Stephen Neale, 1992. Paul Grice and the Philosophy of Language. *Linguistics & Philosophy* 15:5:509–559.
- R. Cranston Paull and Theodore Sider, 1992. In Defense of Global Supervenience. *Philosophy and Phenomenological Research* 52(4):833–854.
- Frank P. Ramsey, 1927. Facts and Propositions. In his *Philosophical Papers*, 34–51. Oxford: Basil Blackwell.
- Jonathan Schaffer, 2009. On What Grounds What. In *Metametaphysics: New Essays on the Foundations of Ontology*, edited by David Manley David Chalmers and Ryan Wasserman. Oxford: Oxford University Press.
- Stephen Schiffer, 1972. *Meaning*. Oxford: Oxford University Press.

Scott Soames, 2005. *Reference and Description: The Case Against Two-Dimensionalism*. Princeton, NJ: Princeton University Press.

Robert Stalnaker, 1996. Varieties of Supervenience. *Philosophical Perspectives* 11:221–241.