Radical Anti-Disquotationalism

Andrew Bacon*

PLEASE NOTE: This is not the final draft

Abstract

A number of 'no-proposition' approaches to the liar paradox find themselves implicitly committed to a moderate disquotational principle: the principle that if an utterance of the sentence 'P' says anything at all, it says that P (with suitable restrictions).[1] I show that this principle alone is responsible for the revenge paradoxes that plague this view. I instead propose a view in which there are several closely related language-world relations playing the 'semantic expressing' role, none of which is more central to semantic theorizing than any other. I use this thesis about language and the negative result about disquotation to motivate the view that people do say things with utterances of paradoxical sentences, although they do not say the proposition you'd always expect, as articulated with a disquotational principle.

Consider a self-referential utterance, u, of the sentence 'u is not true'. According to one widespread and appealing intuition when one makes a semantically paradoxical utterance such as u one simply does not succeed in saying anything. Call this the *no proposition theory*.

No proposition theorists reject the disquotational assumption that utterances of 'u is not true' say that u is not true on the grounds that some utterances of 'u is not true' do not say anything at all. However, they may nonetheless subscribe to a qualified version of the principle that says that if an utterance of 'u is not true' says *anything at all* it says that u is not true. In this paper I shall show that such views are susceptible to a version of the revenge paradoxes. An examination of these paradoxes suggests a view in which paradoxical utterances such as u *do* say things, although they do not say what you might expect them to (in this case, that u is not true). I shall show, moreover, that this phenomena falls out of general considerations about the relation between language and the world, and is much more widespread than many have thought.

In this paper we will be primarily be investigating the corner of philosophical space that accepts classical logic and that admits quantifiers that bind variables taking sentence position. Unless otherwise stated, that framework will be assumed throughout.[2]

1 No Proposition Accounts of the Paradoxes

The most straightforward version of the no proposition theory maintains that sentences like 'u is not true' are completely meaningless, and thus that any attempt to say something by making an utterance of this sentence would fail. Call this the *non-contextual* view. On the non-contextual view, then, all utterances of a paradoxical sentence are equally bad.

*Thanks to ...

[1] Restrictions might exclude sentences which express different propositions in different contexts, such as sentences involving indexical expressions.

[2] Of course, neither of these assumptions are uncontentious; but see Prior [19] and Williamson [28] for a defense of the intelligibility of quantification into sentence position, and Williamson [29] for general methodological remarks about the application of classical logic to the liar paradox. For a recent approach to the liar that relaxes classical propositional logic, see Field [8]. For a discussion of approaches that relax the classical rules for the propositional quantifiers see Bacon, Hawthorne and Uzquiano [2].

However, a central component of many extant no-proposition accounts of the semantic paradoxes is the contention that while one utterance of a sentence might fail to say anything, a distinct utterance of the *very same* sentence might result in something being said — perhaps even something true and important to the theorist.[3] Suppose u is an utterance of the sentence 'u is not true'. If u doesn't say anything because it is paradoxical then, presumably, u is neither true nor false. So, in particular, u is not true.

The important thing to note is that I made this last point, that u is not true, by producing a different token of the sentence 'u is not true'; when I produce a token of 'u is not true' I say something true and theoretically enlightening.

This is how Williamson describes the process which brought about the change between the utterance u and my utterance:

> "We start with one set of correlative meanings for 'say', 'true' and 'false'; we use them to construct a sentence that says nothing in that sense of 'say'; but reflection on that sentence causes normal speakers to give 'say', 'true' and 'false' a new set of correlative meanings, much like the previous ones except that the sentence in question says something in the new sense of 'say'; the process can be repeated indefinitely." [25] p15

According to Williamson consideration of the semantic paradoxes forces us to recognise a more expansive use of the word 'says' (and by extension 'true' and 'false'). Self-referential utterances involving these more expansive uses of these words force us to recognise even more expansive uses, and this process repeats indefinitely. As I find Williamson's way of setting up the issues congenial, I will use his theory as a springboard for my investigation into 'no proposition' accounts of the liar. Germs of this idea trace back at least as far as medieval scholars (the 'cassationists') and have been developed in more recent times by Bar-Hillel [**?**], Whiteley [23], Prior [18], Goldstein ([12]), Gaifman [10], Simmons [20], Weir [22], Rosenkrantz and Sarkohi [**?**] and many others.[4]

Since it is granted that some utterances of 'u is not true' do say something, there is a *prima facie* onus on us to explain why u in particular doesn't say anything. The explanation cannot simply be that u is self-referential, for there are self-referential utterances that do say things; an email beginning 'I'm just sending you this message because...' surely says something even though it's self-referential.

An explanation for why u doesn't say anything is available if we help ourselves to the following principles:

> Utterance Truth: If an utterance says that P then it is true if and only if P
>
> Moderate Disquotation: If an utterance of 'u is not true' says anything at all, it says that u is not true.

Together these two principles entail that u does not say anything. For if u did say something then according to the second principle, since u is an utterance of the sentence 'u is not true', it would say that u was not true. Then by the first principle it would follow that u was true if and only if it wasn't. This is, of course, all perfectly consistent with the possibility that other utterances of 'u is not true' do say things; if v is a distinct utterance of 'u is not true' then all we can derive from the assumption that it says something is that v is true if and only if u isn't true.

The first principle is explicitly endorsed by Williamson.[5] Typically we speak of people, not utterances, as saying things; however it is a fairly harmless abuse of convention to extend the notion to the

[3] See, for example, Williamson [25], Goldstein ([12]) and Gaifman [10] among others.

[4] Although 'no-proposition' theories will be the primary focus of this paper, these views are also closely related to a broader class of contextualist theories which includes Parsons [17], Burge [3], and Glanzberg [11]. Note that Williamson does not quite fit the traditional picture: he argues that the paradoxes produce a shift in the conventional meaning of the word 'says'. Thus both the content and character associated with a sentence can vary from utterance to utterance. The important point, for my purposes, is the variability of the content of tokens of the same sentence.

[5] See principle T on p. 12. The difference is that Williamson talks about sentences in contexts rather than utterances. He notes that this principle is importantly different from the T-schema; indeed Utterance Truth is consistent in classical logic.

utterances that facilitated the saying. A couple of points need to be stressed about this convention, however. Firstly, speech reports are notoriously flexible – I can use the expression 'so and so said that P' to cover things they merely intended to communicate, or simply things they hinted at. I am primarily interested in the slightly more regimented uses of 'says' that is the focus of much semantic theorizing (see Grice [13]).[6] Secondly it is important to note that, even in this regimented sense of 'says', it's not implausible that when people say something they usually also say some of the obvious logical consequences of that thing. If this is right then 'what u was used to say' must be understood rather strictly as something like 'the conjunction of everything u was used to say' or 'the strongest thing u was used to say', or else we should not expect the first principle to hold (see Andjelkovic and Williamson [1] p230-232.) This trick is always possible given the assumption that the propositions form a complete Boolean algebra.[7]

Utterance Truth is related to a slogan of truth conditional semantics: that what a sentence means in a context are the conditions under which it is true. However, in reality it encodes little more than a choice to use the words 'says' and 'true' in a related way when applying them to utterances: that an utterance is true if whatever it says is true. If we permit ourselves a device that allows us to quantify into sentence position, as Williamson does, one can give a definition of utterance truth from utterance saying as follows:

> u is true if and only if: (i) for some P such that u says that P, P, and (ii) for no P such that u say that P, is it not the case that P.[8]

A bit more colloquially, u is true iff it says something true and nothing false. Note that the colloquial way of paraphrasing this definition makes use of the notions of propositional truth and falsity, even though they make no appearance in the above definition: here and elsewhere I shall often use such paraphrases for convenience, but it is important to note that truth and falsity, as they apply to propositions, make no appearance in the sentences they are paraphrasing.

Given the assumption that no utterance says more than one thing, one can prove Utterance Truth from this definition. Indeed, this assumption is already guaranteed by our convention to understand 'says' so that it relates u to P only if P is he conjunction of everything u said in the looser sense. Since *Utterance Truth* is guaranteed by our definitions, it cannot be responsible for the paradox.

The second principle was also needed to explain why u doesn't say anything. As a simple model to demonstrate that no inconsistency follows without it, note that if u had said that snow was white, then u would have said something, and it would be straightforwardly true assuming our definition of utterance truth. Without this moderate form of disquotation there is no logical guarantee that u doesn't say anything, although it could be added as an extra-logical assumption.

It should be stressed that Moderate Disquotation is not a schema, it is just a principle about a particular sentence. Some instances of the more general schema are hardly plausible: if you utter the sentence 'I'm hungry' you've said something, but you haven't said that *I'm* hungry you've said that *you* are hungry. This moderate disquotational principle highlights that, according to this kind of no-proposition theory, the paradoxical phenomenon is not quite the same as the phenomenon displayed by indexicals like 'I' and ordinary context sensitive words like 'tall'. In the latter case two utterances of the sentence 'Fred is tall', for example, can be used to say different things. In the case of 'u is not

[6] Even with all the clarifications and stipulations that philosophers perform before using the word, its important to note that it still retains a substantial amount of flexibility.

[7] Those who prefer to theorise in terms of structured propositions, and who thus deny this assumption, will typically countenance many such conjunctions. Even if there are multiple such conjunctions, they will all be materially equivalent (indeed, logically equivalent) since they only differ over the order of the conjuncts. The paradoxes we will be considering pose just as much a problem for the view that each sentence strictly says several materially equivalent propositions as for the view that each sentence strictly says one thing (cf Prior's Theorem in section 2).

[8] This definition may have to be modified if we want to make it friendly to propositional temporalism: the view that there are propositions that are sometimes true and sometimes false. The natural modification would be: u is true $=_{Def}$ for some P, u said that P, and P was true at the time u was uttered and for any P, if u said that P then at the time u was uttered it was the case that P. I will not make any attempt to control for this complication in what follows.

true' two utterances of this sentence can differ with regard to whether they express a proposition at all but given the moderate form of disquotation above no two utterances of 'u is not true' that express a proposition can express different propositions: if they both express propositions, they must both express the proposition that u is not true.[9]

In this paper I will endeavour to show that, despite appearances, u actually *does* say something, even if it cannot be the thing you'd articulate using an disquotational principle.[10]

My main goal in section [?] will be to show that Moderate Disquotation is untenable: it straightforwardly gives rise to further paradoxes. This undermines our original argument for thinking that u does not say anything. I argue instead for an alternative thoroughly anti-disquotational view. We begin with the observation that verbs like 'says that' are vague, semantically plastic and context sensitive. These considerations naturally lead us to the thesis of Semantic Pluralism: that there are a large number of similar but equally natural language-world relations, with none uniquely playing the honorific role of the 'expressing relation' that takes a central place in semantic theorizing. The view is radical in the following sense: once we have noted that there are several distinct semantically important relations I shall argue that few, if any, of these relations relate even utterances of the (apparently unproblematic) sentence 'snow is white' to the proposition that snow is white.

The rest of the paper will be structured as follows. In section 2 I outline and defend an important theorem in the logic of 'saying that' due to Arthur Prior. In section 3.2 I argue that any theory endorsing certain instances of moderate disquotation are untenable. I go on to describe a view which rejects that form of disquotationalism in 4. In doing so I describe (and show consistent) a formal theory that is completely self-applicable. It can talk about its own semantic apparatus, it can state its own principles of compositionality, and moreover its theorems can be shown to be true.[11]

2 Paradoxes of Indirect Discourse

I have been theorizing with the familiar notion of what a person has said, as captured by indirect speech reports of the form 'so-and-so said that P' (understood in the slightly regimented way I outlined

[9] Note that some contextualists (e.g. Simmons) often seem to speak as if u does express a proposition; u is defective, rather, because the proposition it expresses is neither true nor false. If the proposition itself was defective in some way, then, given the point that all utterances of 'u is not true' that say anything say the same thing, all of the other utterances expressing that proposition would be defective. This would be disastrous since the contextual no-proposition theorist wants to make utterances of the sentence 'utterance u is not true' themselves when they describe their view.

[10] A tentative (but suggestive) argument for this conclusion is already available: the above theory provides an attractive account of contingent liars that would otherwise pose a puzzle for no-proposition accounts. If I am standing in an unknown room and I see the sentence 'the only sentence token written on the board in room 101 is untrue' written on an otherwise blank board, I appear to be in a position to rule out the epistemic possibility that the board in room 101 has nothing written on it but the sentence 'snow is white'. This is a quite general fact about the reaction of English speakers to seeing that string of symbols written on a board, assuming they take it to be a reliable source of truths. The default explanation of this fact would be that the sentence token expresses a proposition inconsistent with this possibility. Indeed one might think that it is facts like these that constitute what it means for a sentence to express a proposition that's inconsistent with that kind of possibility. The theory I sketched corroborates the default explanation whereas the no-proposition theory does not, for if you are in fact standing in room 101 the token would not have expressed a proposition. Even if this is not a decisive reason to think that this sentence token says something, even when you are standing in room 101, there is clearly *some* relation between that token and a proposition, and one would want some story about analogous paradoxes for that relation. For example, you can still talk about the set of worlds that competent English speakers will rule out upon accepting the token as trustworthy, or the proposition the token *would* have expressed if I had wheeled the board into the hallway (so it was no longer located in room 101). For a discussion of the latter case, see Zardini [30] p563-4. Analogous paradoxes involving these relations suggest that simply denying that a proposition is expressed won't get to the heart of the matter.

[11] Let me briefly note that in what follows I will be setting aside a certain kind of contextualist approach to the liar modeled on the idea that propositions are indefinitely extensible much like the way the set theoretic heirarchy is sometimes claimed to be (see Parsons [17] and Glanzberg [11]). According to these theorists, reflection on the paradoxes causes the domain of propositions to expand in such a way that a proposition becomes available for the liar utterance to express. It would be misleading to call these theorists 'no-proposition' theorists, so I shall reserve for them the term 'indefinite extensibilists'. Indefinite extensibilists face a fairly well known obstacle: they have trouble expressing their own view (see Williamson [27] section V).

earlier). From this we can indirectly introduce the notion of what an utterance was used to say — roughly, if a person said that P, and this was facilitated in some essential way by an utterance or sentence token of some other kind, then that utterance can be thought of as 'saying that' P. Once this is understood we can then introduce utterance truth in way described earlier using quantification into sentence position. It is important to note how this notion differs from the more technical notion of a sentence being true in a context and a language. My notion applies to utterances and not sentences, and it is not relative to a language or a context: if a person succeeded in saying that P by making an utterance it matters not what language or context that utterance was made in, that utterance is true just in case P.

My strategy in the following will be to shed light on the semantic paradoxes by figuring out what people have said when they make utterances of paradoxical sentences. There are, of course, variant paradoxes that are not directly treated by this approach — paradoxes involving a more technical piece of philosophical vocabulary: the notion of a sentence 'semantically expressing' a proposition in a context, and the correlated notion of truth in a context. According to the simplest theory of this relation a sentence semantically expresses the proposition that P in a context if and only if an utterance of that sentence made in that context would result in one saying that P. Given the simplest theory, the focus on figuring out what people have said is well placed. That said, it is entirely possible that the informal notion of what is said by a person is too flimsy to do the theoretical jobs that we need the semantic expressing relation for. Nonetheless, I shall adopt the simplest theory as a placeholder in what follows, as I think it is a good enough approximation of the truth. If we can figure out how the semantic paradoxes behave on this theory we are on track to figuring out how they work more generally.

Since indirect speech reports will play an important role in this investigation we will begin by looking at an important limitative theorem concerning their logic due to Arthur Prior (Prior [18] and Prior [19]). It is important in the following to keep in mind the difference between indirect and direct speech reports. 'Alice said that snow is white' is an example of the former — it is formalized using an operator expression, and it states that Alice bears a relation to a certain proposition. 'Alice said 'snow is white'' is an example of the latter, and it merely says that Alice uttered 'snow is white'. Prior's theorem does not, without extra assumptions, imply anything about what sentences can or cannot be uttered.

Prior's theorem is formulated against the background of an expressively weak fragment of higher order logic, known as 'quantified propositional logic', a language containing the propositional connectives that also allows quantification directly into the position occupied by a sentence.[12] By contrast, full higher order logic would allow quantification into the position of any grammatical category. Formally speaking quantified propositional logic adds to the syntax of propositional logic a countable set of propositional variables, $p_i, i \in \omega$, and a special quantifier $\forall$. Each propositional variable is a well formed formula, and in addition to the standard clauses for building up complex well formed formulae we stipulate that if p_i is a propositional variable and ϕ a well formed formula, $\forall p_i \phi$ is well formed. Finally we will be interested in the extension of this language by a unary connective, S, which prefixed to any well formed formula, ϕ, produces a well formed formula $S\phi$. S here will stand in for 'Alice said at t that'. Within this context Prior proves a general theorem concerning arbitrary operators which results in seemingly paradoxical results when taken to be about the particular operator S.

The premises required to derive Prior's theorem are surprisingly minimal.[13] An intuitively complete axiomatization of quantified propositional logic would require a richer system, however to derive Prior's theorem one only needs, in addition to the standard principles of classical propositional logic, the principle of universal instantiation for the propositional quantifiers which allows you to infer '...ϕ...' from '$\forall p$ ...p...'. We can axiomatise this system as follows:

[12] This is sometimes misleadingly called 'propositional quantification', although the use of these quantifiers does not commit you to the existence of propositions which are singular entities and can which can be quantified over using the ordinary first order quantifiers.

[13] Prior doesn't explicitly note that the necessary system is this weak, although it is immediate from inspection of his proof.

CL All substitution instances of classical tautologies and the rule of modus ponens.

PUI $\forall p\phi \rightarrow \phi[\psi/p]$ where ψ is substitutable for p in ϕ.[14]

Within this system he derives the following theorem:

PRIOR'S THEOREM $S\forall p(Sp \rightarrow \neg p) \rightarrow \exists p(Sp \wedge p) \wedge \exists p(Sp \wedge \neg p)$

The proof, in Polish notation, can be found in [18]. I include a translation into more familiar notation in the appendix for those interested. Here I will proceed informally.

Following Prior, we can paraphrase the reasoning behind the theorem in ordinary English using singular quantification over propositions. Suppose that Alice says at t that everything that Alice says at t is untrue. It follows that she said something true: if she hadn't, then everything she said at t was untrue – but that's exactly what she said, so something she said is true after all. But if something she said at t was true then she was simply not speaking truthfully when she said that everything she said at t was untrue – that is to say, she said something untrue too, namely, that everything she said at t is untrue.[15]

Two points are worth belaboring at this point. (i) Prior's paraphrase involves singular quantification into the argument of a predicate such as in '... is not the case', presumably committing us to proposition-like entities which can be said to be or fail to be the case. Prior's theorem, on the other hand, does not involve quantification over proposition-like entities, it involves quantification into the position that a sentence occupies which is entirely compatible with there being no abstract objects (it is even compatible with there being nothing at all.[16]) (ii) Prior's paraphrase makes reference to something resembling a propositional truth predicate (i.e. '... is the case'), and falsity predicate ('... is not the case') whereas PRIOR'S THEOREM employs a device that quantifies directly into sentence position and therefore makes no use of propositional truth or falsity. Despite these differences, I will adopt these paraphrases as well – to do otherwise would be cumbersome and would involve taking liberties with English.[17]

3 Against Moderate Disquotation

Anyone who accepts classical logic and the logic of quantification into sentence position must accept PRIOR'S THEOREM. Given straightforward paraphrases of the propositional quantifiers in PRIOR'S THEOREM into English using singular quantifiers, however, it seems that we must accept the following apparent falsehoods:

1. Necessarily, if anyone says that everything they've ever said is false then they said something else at some point.

[14] Intuitively 'ψ is substitutable for p in ϕ' just means that no free sentential variables in ψ get bound when substituted into ϕ for p.

[15] As Prior notes, the theorem is very general; one can interpret S as any number of attitudes such as 'thinks that', 'fears that', 'remembered that', 'asserts that', 'means that', 'writes that', 'hopes that', 'queries whether', 'wonders whether', 'brings it about that', 'supposes that', 'desires that' and so on. I shall not make any attempt to generalize the observations I make about 'says that' to these other attitudes, although there are some obvious parallels.

[16] In models of propositionally quantified logic the domain of the singular quantifiers and the domain of the propositional quantifiers are independent of one another. One can thus construct models in which sentences involving singular quantifiers like 'there are no objects' and 'there are no propositions' are true, while also having a rich theory involving the propositional quantifiers.

[17] One might worry that the singular paraphrases involving a propositional truth predicate commit you to paradoxes similar to those that arise for sentential truth predicate. This is actually not true: it is fairly easy to show the consistency of the propositional T-schema, $T(\mathit{that}\ \phi) \leftrightarrow \phi$, by modeling propositions simply as truth values, or sets of worlds (formalism: prefixing 'that' to a sentence produces a name for a proposition.) One can prove the consistency of this theory by noting that one can recursively define the denotations of the terms formed out of the 'that' subnective by extending the notion of the complexity of a formula to the 'that' terms in the obvious manner. Of course, one could enrich the theory by adding axioms that make the 'that' subnective behave like more like the quotation subnective and the enriched theory would be inconsistent. However this is only a problem for very fine grained theories of propositions; any view in which propositions are sets of indices of some kind can be shown to be consistent.

2. Necessarily, if Alice says at t that everything she's said at t is false, then she's said at least two things at t (something false and something true.)

These claims, and others like them, appear to contradict obvious platitudes. In this section I will try and diffuse this impression.[18]

3.1 What Alice uttered and what she said

What platitudes do 1 and 2 apparently contradict? One platitude is that Alice could perfectly easily have uttered, at t, the *sentence* 'everything Alice is saying at t is false' and have not simultaneously uttered anything else. This much is absolutely clear and any theory must accomodate this fact. What must be contested, then, is whether this fact contradicts 2. All 2 implies is that if, in uttering this sentence, Alice had said, at t, that everything Alice says at t is false, then she would have said something else at t as well. It is crucial to bear in mind the point emphasized earlier that the connection between the sentence one utters and what one says (if anything) by uttering that sentence, is not straightforward and is certainly not governed by a simple disquotational principle: if John said that he was hungry, for example, he almost certainly wouldn't have achieved this by uttering the sentence 'he was hungry', for he would use this sentence only to ascribe past hungriness to someone else. Yet this is exactly what one would get if you unthinkingly applied a naïve disquotational principle.

Prior's theorem, in combination with this platitude, therefore also places some constraints on this relation that holds between utterances and propositions. Let us suppose that Alice does, in fact, utter the sentence 'everything I am saying at t is false' at t, and nothing else and let us ask what she thereby said. Here are the answers to this question that are consistent with Prior's theorem:[19]

1. Alice didn't say anything at all when she made this utterance. In particular, she didn't say that everything Alice says at t is false.[20]

2. Alice did say that everything she says at t is false, but in doing so she also said something else (and moreover, she said at least two propositions with opposite truth value.)[21]

3. Alice said exactly one thing, but whatever it was that she said, it wasn't that everything she said at t is false.[22]

The first upshot, which can be seen just by surveying these options, is that we cannot apply the naïve disquotational method to determine what Alice said in this case. We must reject:

> DISQUOTATION: Anyone who utters the sentence 'everything Alice says at t is untrue' (with the right intentions, etc) thereby says that everything Alice says at t is untrue and nothing else.

A second striking feature is that according to two of the three options Alice says something completely unexpected. According to the third option Alice says exactly one thing. But since, according to that view, this proposition is not the proposition that everything Alice said at t is false we are left wondering what this mysterious proposition does say. According to the second option Alice says two

[18] One strategy for deflating this objection, which I won't pursue, is to insist that these sentences are not strictly instances of Prior's theorem because they involve singular quantification. One might try to insist that the real instances of Prior's theorem are not counterintuitive because they are not part of ordinary English and are therefore not sentences we have any intuitions about.

[19] There is another option not covered by the below: that Alice says two or more things none of which are the proposition that everything Alice says at t is false. This is not particularly attractive, since it seems to combine the unattractive aspects of the second and third options. Moreover, note that the multiple proposition view (view (2) or its less attractive non-disquotational variant) has been tacitly ruled out by our earlier stipulation that we understand 'says' strictly as the relation holding between a person and the conjunction of things that they've said in the looser sense.

[20] This corresponds to the standard no-proposition line, see the references in section 1.

[21] This option is less well represented in the modern literature, although it was the view of Thomas Bradwardine, and some subsequent medieval scholars.

[22] This option is also not very well represented in the modern literature, but see Nicholas J.J. Smith [**?**].

things, although, at least one of these things is a mysterious proposition. For the time being I shall refer to both of these views as *mysterious proposition* views.

It seems, then, that the no-proposition view (the first option) has a distinct advantage since it seems to avoid these mysterious propositions. In this section I shall argue that this is not the case: the general thought that we don't express an unexpected proposition is untenable. We can state the idea that utterances like Alice's should say the expected proposition, if they say anything at all, more precisely in terms of the following weakened disquotational principle (a variant of the principle MODERATE DISQUOTATION):

> MODERATE DISQUOTATION*: Anyone who utters the sentence 'everything Alice says at t is untrue' and says anything at all by it, says that everything Alice says at t is untrue.[23]

To completely rule out mysterious propositions being expressed one would have to add that the proposition that everything Alice says at t is false is the *only* thing one would end up saying with an utterance of this sentence, if you were to say anything at all. However, I shall actually be refuting the weaker thesis encapsulated by the above principle. Note that, in so far as these questions are discussed at all by contemporary no-proposition theorists, this moderate form of disquotation is often implicitly being assumed, not for context sensitive language in general, but for the kind of context sensitivity that arises with sentences like the one Alice uttered.[24]

The idea that utterances of 'everything Alice says at t is untrue' say what you'd expect, if they say anything at all, is what makes options (1) and (2) more attractive than (3). But if MODERATE DISQUOTATION has failures the third option begins to look more attractive: if you're going to have to countenance mysterious propositions *anyway* you might as well go for the view that deviates least from orthodoxy.

3.2 Moderate disquotation

In section [**?**] we introduced two kinds of no-proposition theory. The non-contextual version was subject to the problem that they have difficulties expressing their own view: if u is a paradoxical utterance of the sentence 'u is not true', then it's part of the non-contextual no-proposition view that u is not true (because it doesn't express a proposition). The natural way to express this is by uttering an instance of the same sentence, 'u is not true'. But on this view all utterances of this sentence fail to say anything. In this section we shall argue that the contextual version of the no-proposition view is also subject to this sort of problem.

Recall that an utterance is true if it says something true on the occasion on which it is made, and nothing it says is false. Here, as before, the apparent appeal to propositional truth and falsity in this definition is eliminable in favour of quantification into sentence position: in Prior's language, u is true iff $\exists P(Say(u, P) \wedge P) \wedge \forall P(Say(u, P) \rightarrow P)$ where $Say(u, P)$ means 'u was used to say that P'.

[23] MODERATE DISQUOTATION* is clearly motivated by the same general sort of intuition that MODERATE DISQUOTATION is. Note that strictly speaking this principle is implausible. For example the quantifier 'everything' is sensitive to a contextually salient domain, so utterances of this sentence where a different domain is salient will result in something else being said. I think it is possible to control for this by introducing a context insensitive quantifier stipulated to range universally over all propositions. The possibility of doing this is not uncontroversial in this context (see Glanzberg [11] and Parsons [17]) but it is an assumption that many, including Williamson, share. Another issue is that presumably the word 'says' is context sensitive in the ordinary sense; I think this worry gets closer to the heart of the problem, and I pursue this in the later sections.

[24] See, for example, Goldstein's paper 'A Consistent Way with Paradox' [12] which makes several implicit appeals to this principle. It is worth contrasting the no-proposition theorists, who typically accept MODERATE DISQUOTATION, with the indefinite extensibilists mentioned earlier, who have the resources to deny it. According to these theorists the word 'true' is short for 'expresses some true proposition', and sentences involve the expression 'some proposition' can express different things in different contexts. Thus, unless we are in Alice's context when we utter MODERATE DISQUOTATION, we could end up expressing a falsehood. Note, finally, that MODERATE DISQUOTATION is vacuously accepted by non-contextual non-proposition views, for such views maintain that no utterance of 'everything Alice says at t is untrue' says anything.

Consider now the sentence L = 'no utterance of L both says something true and nothing false'.[25] Generalising moderate disquotation to this sentence gives us the following:

> If an utterance of 'no utterance of L says something true and nothing false' says anything at all it says that no utterance of L says something true and nothing false.

Thankfully this principle does not collapse into inconsistency: to get an inconsistency you'd have to assume that some utterance of L said truths and only truths. You can get around the paradox by maintaining that there are no utterances of L like this perhaps by insisting that no utterance of L says anything at all. In fact, if you properly formalize the moderate disquotational principle above in terms of propositional quantification, you can rigorously prove from this the result that there are no true utterances of L (utterances which say truths and only truths) in propositionally quantified classical logic:

> THEOREM: From the above instance of moderate disquotation one can prove (in propositionally quantified logic) that no utterance of L says something true and nothing false.

See the last appendix of section 5.

Let me give you the gist of this argument informally:

1. **(Premise)** If an utterance of 'no utterance of L says something true and nothing false' says anything at all, it says that no utterance of L says something true and nothing false. (The premise is just an instance of MODERATE DISQUOTATION.)

2. If an utterance of L says something true and nothing false, it says something.

3. If an utterance of L says something, it says that no utterance of L says something true and nothing false. (from 1, the fact that L='no utterance of L says something true and nothing false' and Leibniz's law.)

4. If an utterance of L says nothing false and says that no utterance of L says something true and nothing false, then no utterance of L says something true and nothing false.

5. So, if an utterance of L says something true and nothing false, then no utterance of L says something true and nothing false. (from 2-4)

6. No utterance of L says something true and nothing false. (From 5.)

Assuming classical logic and the classical logic of the propositional quantifiers, the inference to (6) with (1) as the only premise can be formalized and proven as in appendix 5. Note in particular that all invocations of propositional truth and falsity ('says a truth', 'says nothing false', etc) can be eliminated in favour of quantification into sentence position.

So far so good: if we accept moderate disquotation we know that no utterance of L is true (i.e. says something true and nothing false.) (6) is an interesting theorem, and anybody who endorses the view under consideration should go about asserting (6). But there is a problem: theorem (6), just is the sentence L itself! Thus anyone who tries to utter (6) will have made an utterance which isn't true: an utterance of (6) just is an utterance of L, and as we just established, no utterance of L is true in the sense that it says truths and no falsehoods. The situation is just as dire as with the non-contextualist

[25] This is related to a fairly standard revenge problem for no-proposition token-based theories. Hazen [15], for example, considers the sentence L='Every sentence token equiform with L is false' (utterance falsehood has been substituted for something more explicit in my version). Williamson discusses the variant L = 'L is not true in any context'. The following is an attempt to explicitly formalise these arguments and show that they in fact have exactly one premise (other than propositionally quantified classical logic): an instance of moderate disquotation.

no-proposition view – when the contextualist no-proposition view attempts to describe her view she either fails to say anything with her utterances or she says falsehoods.[26]

The reader should note that unlike the unrestricted principle, DISQUOTATION, the moderate version is not inconsistent with the assumption that certain utterances have been made, although endorsing it involves believing that no utterance of L is true. I have therefore, of course, overlooked a coherent position that involves believing that no utterance of L is true but refraining from ever uttering the sentence which appears to express this belief. I think it is an unattractive view, not because it is inconsistent, but because it is hard to communicate it – people can certainly believe it, but they can't express their beliefs by uttering (6), or things that seem to entail (6) like 'no utterance of L says anything' or 'every utterance of (6) says at least one falsehood'.

So far we have just been discussing the no-proposition view. What of the multiple proposition view (i.e. option two): the view that Alice says two things, one of which is the proposition that everything Alice says at t is untrue? To make sense of this view we must be understanding 'says' in the loose rather than the strict sense.

Note, firstly, that this view doesn't avoid the mysterious proposition – there is still at least one other proposition that Alice says at t, and it's not the proposition that everything Alice said at t is untrue. More importantly, note that the view still seems to accord with MODERATE DISQUOTATION. In this case the restriction to utterances that say at least one thing becomes vacuous and we get to say something stronger:

> Anyone who utters the sentence 'everything Alice says at t is untrue' thereby says that everything Alice says at t is untrue (possibly along with other things.)

This might at first seem like a point in its favour, as moderate disquotation admittedly has a great deal of intuitive pull. Of course, it is also its undoing since the above paradoxes we have been discussing apply to any view that endorses moderate disquotation. In this case we can infer, as before, that no utterance of L is true in the sense that it is used to say only true propositions. In this case, although every utterance of L will say something, all utterances of L will say at least one falsehood (and possible also a truth in some cases.) This seems just as bad, since in order for the theorist to report this theorem she will have to utter the sentence 'all utterances of L are untrue', and will have thereby said a falsehood.

One might, at this juncture, argue that the multiple proposition view isn't in as dire a dilemma as the no proposition view: although we ultimately end up saying falsehoods with utterances of (6), we also say the important truth that everything Alice says at t is untrue. We might think that this is the real truth that we intended to say by making this utterance, and that we represent ourselves as knowing when we make the utterance. This strategy is not actually limited to the multiple-proposition view: the no-proposition view could also maintain that although we fail to say anything with utterances of (6), it's clear what we *intended* to say.[27] When I utter (6) I represent myself as knowing that no utterance of L says something true and nothing false.

It is important at this stage not to forget the generality of Prior's theorem. Prior's theorem applies to anything that's formally an operator, and thus can be applied to the attitudes 'S represents herself

[26] It should be noted that strictly speaking quantifier phrases like 'no utterance' are context sensitive in the ordinary sense, so one has to be a little bit careful how one formulates these instances of MODERATE DISQUOTATION. One could, for example, artificially introduce an unrestricted quantifier into the language which, in every context, allows one to quantify unrestrictedly over all utterances. Note that this extremely mild constraint certainly doesn't commit us to anything as controversial as a completely unrestricted quantifier or a universal domain. The class of utterances in our universe is plausibly finite. Unlike certain large classes, nobody would seriously deny that this class forms a set. A version of the indefinite extensibilist idea – that in any context one can always find a context in which more sets exist – doesn't really extend to concrete entities such as utterances: these things can't just pop into existence with a more expansive use of our quantifiers. It is a straightforwardly empirical matter which and how many utterances there are.

More worrying is the potential context sensitivity of the propositional quantifiers. Indefinite extensibilists insist that propositions are indefinitely extensible in the way sets are, so that it is not possible to introduce a context insensitive quantifier quantifying unrestrictedly over all propositions. Such a view has the means to resist our argument, although they fall afoul of other expressive problems (I shan't rehash these problems here, but see Williamson [27] section V).

[27] Or, perhaps, what is *saliently* said.

as knowing that' and 'S intended to communicate that' so parallel arguments involving these operators show that the multiple proposition theorist and the non-proposition theorist cannot even intend to communicate, or represent herself as knowing only truths by uttering the variant of (6) that they are committed to in these parallel arguments.

3.3 A Williamsonian rejoinder?

Although the conclusion of the last section poses a trouble for many contextualist views, Williamson makes a distinctive further claim about the semantic paradoxes that might in principle help make this conclusion sound better. According to Williamson 'reflection on [the paradoxical] sentence causes normal speakers to give 'say', 'true' and 'false' a new set of correlative meanings, much like the previous ones except that the sentence in question says something in the new sense of 'say".

Here is why this might help. Suppose that Bob is about to reflect upon the sentence L, and conclude that no utterance of L is true by making an utterance of L itself. As we pointed out Bob will not succeed in saying anything. However, according to Williamson, once Bob has gone through this type of reasoning the words 'says', 'true' and 'false' shift their meanings.[28] For the sake of argument suppose they now denote the relation of saying$_+$ and the properties of truth$_+$ and falsity$_+$. So while Bob doesn't succeed in saying anything with his utterance, perhaps he does succeed in saying$_+$ something, and although his utterance isn't true, it may well be true$_+$. Note also that even though he doesn't say anything and his utterance is not true, after the change has occurred we can truthfully report these facts using sentences like 'Bob said something' and 'Bob's utterance was true', in the sense that these utterances will be true and true$^+$ at times after the change has occurred.[29]

So although the Williamsonian can't express their view that no utterance of L says anything using the sentence 'no utterance of L says anything', they are in a position to express$^+$ their view using this sentence. Why should we care about this result? Well, provided expressing$^+$ and cognate notions play roughly the same role in the communication of beliefs among language speakers as expressing and cognate notions did, then the fact that we can express$^+$ the view seems like a perfectly good response to the objection that no-proposition theorists can't communicate their view. This response therefore rests (although this is not something that Williamson emphasizes) on an instance of the thesis I called Semantic Pluralism earlier: that there is more than one natural language-world relation playing the 'expressing' role that takes a central place in semantic theorizing.

It will be instructive to consider a toy model of Williamson's account. I make no claim that the following represents Williamson's views accurately or completely – indeed there are several natural choice points. I just want a fleshed out picture on the table so that we have something relatively concrete for contrast.

Assume there are three days: day 1, day 2 and day 3 during which exactly one shift in the meaning of the word 'says' occurs per day:

> On day 1, 'says' has a certain meaning, the relation of saying$_1$, on day 2 it means saying$_2$, and on day$_3$ it means saying$_3$.

The above description, as I have stated it, rests on an assumption that must be justified. For it is a non-trivial assumption, in the present setting, that we can talk unambiguously about *the meaning* of a substential expression like 'says' (for example, when we said "says' means the relation of saying$_1$ on day 1'). Presumably what 'says' means on day 1 is closely related to what utterances involving that word

[28] This feature of Williamson's approach distinguishes him from other no-proposition accounts.

[29] For the purposes of this discussion I have supressed some features of of Williamson's account. For example, Williamson is using the word 'says' as a relation between sentences, contexts, languages and propositions; the expression 'semantically expresses' would be a more apt name for this notion. I am using the word in its ordinary sense, as it appears in speech reports like 'so and so said that P', which is not a language or context relative notion; the notion of utterance truth and utterance saying are similarly not language relative, as they are defined in terms of this notion. For Williamson reflection upon the semantic paradoxes brings with it a change of language. Note that this makes no difference as far as my language-independent notions of saying and truth are concerned: whatever language Alice is speaking, her utterance is true if for some P she said that P in virtue of having made that utterance, and P.

say on day 1. And just as there are multiple saying relations, relating utterances of complete sentences to propositions, there are potentially multiple corresponding subsentential meaning relations, so we can talk about what the word 'says' means$_1$, means$_2$ and means$_3$ on each of the days. This therefore muddies the relation that the word 'says' has to the relation says$_1$ on day 1, says$_2$ on day 2 and says$_3$ on day 3. Could the word 'says', for instance, mean$_1$ the saying$_1$ relation on day one, whilst simultaneously meaning$_2$ some other relation? To avoid the dizzying effects of too much semantic pluralism, let us assume that the meaning$_1$, meaning$_2$ and meaning$_3$ of 'says' always agree on any given day (despite differing from day to day.)

This apparently conflicts with the fact that what is said$_1$, said$_2$ and said$_3$ by utterances of sentences containing the word 'says' will not be the same on a given day. However, this is because the different saying relations differ about whether the utterance in question says anything at all. Thus we may assume that the differences between what is said$_1$ and said$_2$ (say) are not due to differences in the subsentential meaning$_1$ and meaning$_2$ of 'says', but due to a failure of compositionality: an utterance on day 1 might say$_2$ something without saying$_1$ anything, even when their subsential parts have the same meanings$_1$ as meanings$_2$ on day 1.

Let us also assume that Williamson's diagnosis of the paradoxes is true$_i$ ($i = 1, 2, 3$) whichever day he makes the diagnosis: utterances of "says' changes its meaning between day 1 and 2 and between day 2 and 3' is true$_i$ ($i = 1, 2, 3$) on days 1, 2 and 3. Thus the word 'says' means$_1$ three different things over the period of three days, it also means$_2$ different things on each day, and also means$_3$ different things on each day. Given our previous remarks these will be the same three things: saying$_1$, saying$_2$ and saying$_3$.

So to summarize: on each day the word 'says' means$_1$, means$_2$ and means$_3$ the same thing. Although the meaning$_1$, meaning$_2$ and meaning$_3$ changes from day to day, they do so in unison. It is not true, however, that utterances of 'every utterance of L is untrue' say$_1$, say$_2$ and say$_3$ the same thing on a given day: on day 1, for example utterances of this sentence don't say$_1$ anything, even though they say$_2$ something. Even though their subsentential parts have the same meaning$_1$/meaning$_2$, compositionality$_1$ fails for this particular utterance because it was made on day 1.

We have three saying relations so it is natural to wonder what utterances of the sentence L (='no utterance of L is true') say$_1$, say$_2$ and say$_3$ on the three days respectively. Lets start with saying$_1$. The intuitive picture, I take it, is that on day 1 utterances of L say$_1$ nothing: given the meaning$_1$ of 'says' on day 1, if L were to say$_1$ anything at all it would say$_1$ that no utterance of L is true$_1$, which it cannot on pain of paradox (given the connection between saying$_1$ and truth$_1$.) Then the meaning$_1$ of the word 'says' (and 'true' and 'false') changes and comes to mean$_1$ saying$_2$. Thus on day 2 utterances of L say$_1$ that no utterances of L are true$_2$. Then, on day 3, another change in meaning$_1$ occurs and L comes to say that no utterances of L are true$_3$.

Let us turn to what is said$_2$ on the three different days. Two of the cases should be completely parallel to day 1: L doesn't says$_2$ anything on day 2, and on day 3 a shift of meaning$_2$ occurs and L says$_2$ that no utterance of L is true$_3$. But what do utterances of L say$_2$ on day 1? Presumably they say$_2$ that no utterance of L is true$_1$ on day 1 because the meaning$_2$ of 'true' on day 1 is truth$_1$, and no paradox ensues from assuming that compositionality applies and that L expresses something. It is interesting to note that it follows that neither says$_1$ nor says$_2$ is more expansive than the other: says$_1$ crashes with utterances of L made on day 1, but not day 2, and says$_2$ crashes for utterances of L on day 2 but not day 1. What we say$_3$ on the different days should be similar: we'll end up saying$_3$ things on days 1 and 2 with utterances of L, but we won't say$_3$ anything on day 3.

Note that in the above model, utterances of L made on day 1 don't say$_1$ anything, although they do say$_2$ something. However the reverse is also true: utterances of L made on day 2 don't say$_2$ anything although they do say$_1$ something. Although the former fact accords with Williamson's claim that the meaning of 'says' gets more and more expansive as the days move on, the latter fact contradicts it: saying$_1$ assigns a proposition to an utterance where saying$_2$ doesn't assign anything. One could, of course, artificially modify the model so that it accords with the idea that saying$_2$ is strictly more expansive than saying$_1$, but it seems like there is no good reason to do this. No inconsistency arises

from supposing that utterances of L on day 2 express$_1$ a proposition.

Let us summarize the features of the view I've just sketched. Firstly, the temporal element of the view isn't as important as it might seem at first: on day 1, when we make ordinary utterances we simultaneously say$_1$, say$_2$ and say$_3$ something (there will be some gaps regarding what we say$_1$ with paradoxical utterances, but this is the general picture). However we suggested, in response to the expressive worry raised in the last section, that the three different saying relations are all equally important for describing how we communicate; it's not OK to say$_2$ false things on day 1 any more than it is to say$_1$ false things. The view therefore seems to be one in which utterances on any given day have three important semantic properties; the three semantic statuses of these utterances do not vary from day to day.

Secondly, when the temporal element of the view does come in to play, it has some fairly counterintuitive consequences. For example, it requires a quite remarkable and somewhat mysterious connection between the meaning of certain words and the practices of a certain group of people consisting mostly of logicians and philosophers (the group of people responsible for doing the most reflecting upon paradoxical sentences.) Puzzling questions quickly present themselves: if two people reflect upon a paradoxical sentence in quick succession, the language will successively undergo two quick changes of meaning. However if they reflect upon the sentence at exactly the same time, will the language undergo two changes or one? One might also wonder whether merely thinking about the paradoxes induces the change, or must it be aloud? If it is merely thinking then would it be possible for a non-English speaker to change the meaning of English words just by thinking about the paradoxes? What if, for example, an alien was aware that Alice had attempted to say that everything she's saying at t is false, whilst being completely unaware of how she tried to say it and of the English language quite generally. Could such a person really have an effect on the meaning of English words?

Thirdly, and lastly the view I have described isn't compositional – this is at least a *prima facie* cost. Indeed, if our explanation of our failing to say anything with an utterance boils down to a failure of compositionality, a sort of revenge paradox arises quite straightforwardly. Let the *compositional closure* of some subsentential meaning relations, written 'says^{cc}', be defined as follows:

> An utterance of S says^{cc} that P iff the meanings of the parts of S, when applied to each other in accordance with the syntactic make up of S, compose to the proposition that P.

Even if an utterance fails to say something, provided its subsentential parts having meanings, it will still bear the compositional closure of saying, saying^{cc}, to some proposition. Instances of Prior's theorem can be stated that instead use the saying^{cc} relation.

Indeed, these sorts of reflections have lead me to think that the best way to avoid sentential disquotational principles is to deny disquotational principles as they apply at the subsentential level (such as 'says' means says, 'true' means true, and so on). Since one cannot derive via compositionality the sentential disquotational principles without assuming the substential counterparts, such views have no need to deny compositionality. Let us then turn to my favoured version of this view.

4 Radical Anti-Disquotationalism

The above considerations cast an unfavourable light on the no-proposition and multiple-proposition interpretations of Prior's theorem – that Alice said nothing in uttering her sentence or that she said several things: both positions have trouble expressing their views. This leaves only the third option: That Alice says exactly one thing when she utters the sentence 'everything Alice is saying at t is untrue', but it is not the proposition that everything Alice is saying at t is untrue. This response, however, immediately raises the question: *what on earth* is the proposition that she said in that circumstance, if not the proposition that everything she said at t is untrue? It's all very well to make claims about what she didn't say, but one ought to be able to say something informative about what she did say otherwise it begins to look very mysterious.

In order to explain this puzzle, and much more, I will defend and appeal to the following thesis:

Semantic Pluralism: There is a large number of language-world (and mind-world) relations that are all extremely similar to one another, each one playing similar roles in the theory of communication, and none more deserving of the scrutiny of semantic theorizing than any other.[30]

By a language-world (mind-world relation) I simply mean a relation holding between utterances (or mental tokens) and propositions. The kind of relation you might report as standing between an utterance and a proposition P when the producer of the utterance has said that P, or a belief (or any other attitude) and a proposition P if it is a belief that P. I shall focus on the linguistic case in what follows. In light of the discussion in section [**?**] it is natural to view Williamson's view as a version of semantic pluralism.

I claim that although semantic pluralism is quite a natural view it entails the following much more surprising thesis:

Radical Anti-Disquotationalism: Hardly any of these language-world relations relate utterances of 'snow is white' to the proposition that snow is white.

In fact, I take Radical Anti-Disquotationalism (RAD, for short) to entail a long conjunction of surprising theses like the one above. In general one must resist the urge to narrow down the relations that play the expressing role to those which 'behave disquotationally' (I put this expression in scare quotes because it will soon become clear that 'behaving disquotationally' is not even a well defined property).

I choose the sentence 'snow is white' because it is often taken to be a paradigm case of a sentence that satisfies the disquotational schema. Actually the sentence 'snow is white', like most English sentences, is vague and therefore, at least according to one orthodoxy, context sensitive (although see Cappelen and Lepore [4].) Since context sensitive sentences do not straightforwardly satisfy disquotational principles, and vagueness is so pervasive, we might already find ourselves to be committed to pervasive failures of disquotation. However context sensitive sentences still satisfy qualified versions of disquotational principles — for example, one might insist that utterances of 'snow is white' made in the present context say that snow is white. My arguments for RAD will establish that even these qualified disquotational principles have failures.

Why be interested in radical anti-disquotationalism? While it may be interesting and surprising in its own right, it also encodes exactly the kind of assumption that was responsible for the paradoxes we have been discussing. Once we've accepted a natural picture in which disquotational principles are in general false and fail in paradigm cases of ordinary semantic theorizing, we can see that the semantic paradoxes are actually quite unexceptional and require no revision to our semantic practices. This is shown more rigorously in section 5 where a model of semantic pluralism is given of a language which can reason about its own semantics in a completely general way.

It should be stressed that although RAD encodes an important sense in which the present view is 'anti-disquotational', there are other kinds of disquotational ideas one might want to give voice to. In particular, RAD does not entail that it's always bad to make utterances of disquotational *sentences*. Indeed it's consistent that utterances of the sentence ' 'snow is white' means that snow is white' are related by these language world relations only to truths. We discuss this point further in section 4.4.

4.1 Semantic Pluralism

The business of semantics is to match up bits of language to bits of the world; names to objects, predicates to properties, sentences to propositions and so on. If this was all that was required, matchings of this sort would not be hard to come by – there are lots of uninteresting ways of matching up utterances to propositions, for example. To fall within the purview of semantics a relation between an utterance and a proposition must additionally play some role explaining how that utterance causes

[30] Semantic pluralism is not the same as alethic pluralism, the view that different subject matters require different truth predicates.

beliefs in that proposition amongst speakers of the relevant language. Semantic pluralism is simply the view that there are many relations of this sort that fit the bill.

It would, I think, be incredibly surprising if there was exactly one relation between utterances and propositions that satisfied the kind of communicative role (described in more detail below.) Thus semantic pluralism should have a considerable amount of prima facie appeal. However, it is worth examining the arguments for semantic pluralism in more detail to get a view of the mechanisms by which it comes about.

The central theoretical term for philosophers and semanticists concerned with foundational issues within the theory of meaning is that of a sentence *semantically expressing* a proposition in a language relative to a context. This is a technical notion, although its can be elucidated by its relation to other non-technical notions. Although there is much to be said, the following remarks are at least a start in that direction:

1. We clearly use sentences to communicate with one another, yet what a sentence semantically expresses in a context will in general be much weaker than what is communicated by a given use of that sentence in that context. When I utter a sentence my audience learns that I have vocal chords, a British accent, that I believe what I said, as well as any conversational or conventional implicatures my utterance might carry. None of these is part of what the sentence semantically expresses in that context.

2. What a sentence semantically expresses in a context it often coincides with *what the speaker said* when a speaker makes an utterance of that sentence in that context.

3. If a speaker said that P with an utterance of the sentence S, and it is moreover the case that P, then that utterance is true; in these cases one can talk of the *sentence* being true in that context in the language in question. Thus given 2, we can say that a sentence S is true in L at c if and only if for some P, S semantically expresses P in L at c, and P.[31]

4. Assertion (or assertive utterance) is normatively constrained by the semantic notion of truth: one shouldn't assertively utter a sentence if it is not true.

5. The relation of semantic expressing is governed by a principle of compositionality; for example a disjunction expresses the disunction of the propositions each disjunct expresses. One can generalise this to the subsentential principles with a generalised notion of semantic value.

These remarks, I take it, encapsulate the most important aspects of the role that technical notions like 'semantic value', 'semantic expression', 'truth' and so on, play in philosophical discussions of semantics and in linguistics.[32]

Earlier we made the simplifying assumption that technical talk of a sentence 'semantically expressing' a proposition P in a context can be eliminated in favour of just talking about what an agent uttering that sentence in that context would have said. Although this leaves talk of 'semantic expression' as loose and context sensitive as the notion of 'what is said', this is not entirely inappropriate: it would be remarkable if philosophers had managed to precisely and context insensitively latch onto a singly important relation in the course of their theorizing with phrases like 'semantically expresses'.[33] However the crucial point, which everyone can accept, is the intimate connection the technical notions have have with the pretheoretic notion of *what is said* by an utterance of a sentence.

[31] Earlier in the paper I defined utterance truth as 'saying something true and nothing false'. However if we understand saying and expressing in the strict sense it follows that things say and express no more than one thing. Thus we can drop the clause 'and nothing false' from our definition.

[32] See Soames [21] chapter 4 for a more detailed description of this role.

[33] The eliminativist strategy is not quite so straightforward for the generalized notion of semantic value which applies to semantic categories other than sentences. I suspect we can make good sense of what a person is referring to with a use of a singular term (not to be understood as what they are intending to refer to.) However the slightly more technical notions seem to be needed if we are to extend talk of semantic value to predicates and other subsentential categories.

Let me begin with a suggestive, although not decisive motivation for semantic pluralism. As noted already, speech reports of the form 'so-and-so said that P' are notoriously context sensitive. If Bob, for example, says that he will be up at 5am tomorrow then in some contexts I can report him by saying 'Bob said that he would be up early' (say, if we want to know if he'll be up before noon), whereas in others I cannot, and might even be able to say 'Bob did not say that he would be up early' (for example, if we were in a court of law determining whether Bob had committed purjury.) Thus there is a plethora of distinct but closely related relations that the word 'says' can be used to express, and for each of these conceptions of what it is that a person said with a given utterance there is a corresponding semantic relation relating that utterance to that proposition. Whence, semantic pluralism.[34]

Now the above example employed a use of the word 'says' that was fairly loose – it related Bob to the proposition that Bob would be up early when usually we would describe him as having said something much stronger. Presumably the notion of 'what is said' that is employed in elucidating the notion of semantic expression is a bit more strict than that. However once we have conceded that the word 'says' can be used to mean lots of different things, many of which differ from one another only by tiny factors, it seems unlikely that even philosophers are able to settle on a perfectly consistent and unique use of the word 'says'. More importantly, even if they could, by some tremendous feat, do this, it seems unlikely that they would be picking up an a uniquely important attitude. Many of the relations that the word 'says' expresses can feature in explanations of how people communicate – one can say things like 'A did this and that because B said so and so' in a number of different contexts, and explain A actions in terms of a number of different attitudes B holds towards a certain proposition.

While I take the argument from context sensitivity to be suggestive, it is far from watertight. In most contexts the word 'says' appears to relate people to more than one proposition at a time. It is thus consistent with everything I've said to have a picture in which each utterance is associated with a strongest proposition corresponding to the conjunction of the propositions said, along with a few of its obvious logical consequences.[35] One could insist that all context sensitivity in the word 'says' is context sensitivity deriving from how many consequences of that proposition we also count as having also been said. This view is consistent with the idea that there is a single proposition that is what an utterance has *strictly speaking* been used to say, and which all uses of the word 'says' relate that utterance to, and that any context sensitivity concerns only which of the logical consequences of that proposition come along for the ride. On this picture one could simply identify the proposition semantically expressed with the strongest base proposition that is constant between all uses of the word 'says'; such a view does not obviously motivate semantic pluralism.[36]

A better model of semantic pluralism is suggested by consideration of the vagueness of words like 'says' and 'semantically expresses'.[37] As many people have observed, vague words like 'tall', 'red', 'bald' and so on, typically also exhibit a degree of context sensitivity which is distinctive in that the differences in meaning are only very slight. Moreover, the cloud of propositions that's associated with a sentence like 'Harry is bald' is typically such that there isn't a member of that cloud which entails all the other members of that cloud: there is no proposition in the cloud that could be a plausible candidate for what has strictly speaking been said.

Note, however, that even if vague words are not always context sensitive, it is surely true that when a term is vague there is also a plethora of of closely related denotations of the appropriate semantic type associated with that word, each of which playing roughly the same conceptual role as any other (see Dorr and Hawthorne [7].) For example even if we assume that the predicate 'tall' picks out exactly one property, there are plenty of closely related properties, differing only with regards to

[34] If we adopt the eliminativist strategy suggested above, this conclusion follows directly from the context sensitivity of 'says'.

[35] There is of course still the strict sense of 'says' that you can define as the conjunction of things said in the loose sense. However the data supporting the context sensitivity of 'says' all concern the ordinary use of 'says'.

[36] Thanks to [REF] for pressing me to think more about this type of view.

[37] The following discussion is heavily inspired by the discussion in Dorr and Hawthorne [7].

the minimum heights they are compatible with, each of which seem equally good candidates to be associated with the word 'tall'. The word 'says' is surely the same in this regard. Although we do not have as clear a dimension as height to make the point, surely whether a person has said that P, or whether an utterance semantically expresses P, supervenes on the microphysical facts in some way, thus there will be a host of candidate relations that differ only slightly with regard to which microphysical states they are compatible with.[38]

Before I give some more concrete examples of such relations, let us pause to stress how general these points are. In the above arguments I have primarily focused on the vagueness and context sensitivity of the word 'says'. Perhaps you think the technical notion of semantic expression should not be as closely tied to this pretheoretic notion of 'what is said' as I have assumed. Other conceptions of the central language-world relation are available; perhaps one can spell it out in terms of conventions of truth and trust amongst the linguistic community (Lewis [16]), or in terms of interpretations that maximise the number of true beliefs of the linguistic community (Davidson [5]) or in some other way. Although these accounts place more emphasis on the attitudes of belief and desire and less on the attitude of saying, the considerations we have been considering extend fairly naturally to these other attitudes. However we spell it out, it would be utterly wild to think that the resulting account will not involve words that are vague or context sensitive, and thus it seems incredibly natural to think that there will be more than one relation which plays the semantic expressing role according to these theories. Thus analogous conclusions can be drawn even under the assumption of these alternative pictures.

Let us now make the above ideas more explicit. Here is one way. It is a platitude that as languages evolve words change their meanings – this much is true whatever we take meaning to be. Let us suppose that an expression E now has a different meaning than it did fifty years ago. Now whether this change happened gradually, with the expression continually changing its meaning by small amounts, or in a small number of big jumps, there must have been a moment of time at which the expression had a meaning which it didn't have one nano-second earlier. This change is naturally going to be a function of how the use of the expression is changing over time. Now presumably there is a very similar function of the use of E, calling this 'meaning*', according to which the change in meaning* is delayed by a nano-second relative to the change in meaning. Now, I maintain that it would quite remarkable if the thing I've presumptuously called 'meaning', and not the thing I've called 'meaning*', played a more important role in the theory of communication. For even if we focused exclusively on communication occurring at the nano-second at which they differ, it is hard to see how any important feature of communication could feature one of these relations more centrally than the other.[39]

Here is another reason to think there is a cluster of closely related language-world relations, this time stemming from the observation that there are typically a cluster of closely related propositions associated with each *non*-semantic sentence. For example, in addition to the proposition an utterance of 'snow is white' actually expresses, call that P, there are a cluster of closely related propositions that differ from one another only slightly. Perhaps the strongest thing P entails about the reflectiveness of snow is that snow reflects a particular range of the visible spectrum R, whereas the strongest thing entailed by another proposition in the cluster, Q, is that snow reflects a very slightly shifted range R'. Since, by assumption, neither the range R nor R' is contained within the other, neither P nor Q is entailed by the other (the cluster of propositions associated with 'snow is white' is not linearly ordered by entailment.) Now in addition to the relation that I've (again, presumptuously) called 'semantic expressing' that relates this utterance of 'snow is white' to P, there is another very similar relation that relates utterances of sentences involving the word 'white' to ever so slightly different propositions,

[38] See Dorr and Hawthorne [7] for a way of make the notion of a 'slight difference in microphysical facts' precise. Dorr and Hawthorne make a very similar argument to the the one above, arguing that "In almost all cases, an expressions actual meaning is surrounded by a vast cloud of slight variants which seem just as well qualified to be possible meanings."

[39] Note that even if we model the meaning of a sentence as a cluster of propositions rather than a single one, it's just as clear that a sentence can change the cluster of propositions it means over a period of time and an analogous argument can be run on that assumption.

and in particular relations that match this utterance to Q. Once again it would be remarkable if one of these two relations was theoretically more significant than the other, for the way in which beliefs that P pattern with the use of the sentence 'snow is white' is roughly the same as the way in which beliefs that Q pattern with this usage.

Let me end, then, by making two further comments about the generality of this discussion. Firstly, it is hard to resist the arguments above by maintaining that the semantic expression relation relates each utterance not to a single proposition, but to the cluster of propositions itself. For the clusters themselves can come in clusters of similar clusters of propositions, differing only slightly over what they jointly rule out. In what follows I shall continue to assume that each of the semantic relations we are concerned with is *strict*: they relate each sentence to a unique proposition.

Secondly – although it is beyond the scope of this paper to explore this idea thoroughly – very similar arguments can be run for to mind-world relations by noting that attitude reports are also generally vague and context-sensitive.[40] For example, let b the be mentalese equivalent of the English sentence 'snow is white'. Suppose, by analogy with the linguistic case, that b is a belief that P, where P is a proposition that is inconsistent with certain hypotheses about the reflectivity of snow, and Q is a very similar proposition that is inconsistent with a slightly different range of hypothesis. A parallel with our earlier discussion arises, for there is another belief-like attitude belief*, in which b counts as a belief* that Q, which plays roughly the same role in explaining our actions and verbal behaviour as the attitude of belief did.

Here is some useful terminology for what follows. I shall use 'says_1', 'says_2', 'says_3' as place-holders for a complete spelling out of the various different utterance-proposition relations. The subscripts are not to be thought of as an argument place for a number, they are merely typographical ways of distinguishing the different relations. Corresponding to each of these relations are subsentential semantic relations: for singular terms I shall use 'refers_1', 'refers_2' and so on, and more generally 'denotes_1', 'denotes_2', ... for other semantic categories. Since each family of denotation relations for a given index is supposed to satisfy the role spelled out in (1)-(5) we shall assume that what an utterance says_1 is a function of what its parts denote_1 on that occasion of use (and similarly for what that utterances says_2, says_3 and so on.)

4.2 Semantic Pluralism and Mysterious Propositions

Let us now relate these ideas to our earlier conclusions. We argued that some utterances of L say something, but they do not say what they are intuitively supposed to say. Indeed in section 3 we proved a theorem to the effect that attempting to deny this verbally would involve either saying nothing at all, or saying something false. Moderate disquotation in some sense captures the idea that sentences say what they are supposed to or nothing at all, and thus moderate disquotation has to go. An alternative view maintains that all the paradoxical utterances say something, but what they say is not governed by the moderate disquotational principle; thus, for example, Alice said something with her original utterance of 'everything Alice is saying at t is untrue' but whatever it was she said, it wasn't that everything Alice said at t is untrue.

But if this wasn't what Alice said, what on earth is this mysterious proposition that she *did* she say? It is here that semantic pluralism helps. According to semantic pluralism, the word 'says' is associated with a bunch of distinct but equally important relations and so we have a whole bunch of candidate answers to the mysterious proposition worry. Perhaps, after Alice makes her utterance, she has said at t that everything she said* at t is untrue, where 'saying*' is another semantic relation that plays the saying-role (but is for most purposes very similar to the saying relation).[41]

[40] Many people have noted the vagueness and context sensitivity of belief reports. Although some of these considerations stem from fairly specific considerations of Frege puzzles, others are quite general (see, for example, Field on mental representation [9].)

[41] Note that this idea generalizes to the mental versions of Prior's paradox. Prior's theorem entails, for example, that it's impossible to uniquely think that everything you are thinking is false. On my view when one produces a mentalese token that corresponds to the sentence 'everything I'm thinking is false', I do not think that everything I'm thinking is

Notice, however, that given Semantic Pluralism the above discussion is overly simplistic: we also have a bunch of candidate questions too! It's not clear what one is asking when one asks the question 'what did Alice say?': there's what she said_1, what she said_2 and so on. Further instances of Prior's theorem allow us to get some idea of the situation (letting $i = 1, 2, 3, ...$)

> If Alice is saying_i at t that everything she is saying_i at t is untrue then at t she is saying_i something true and something untrue.

Applying the same line of reasoning to each of the saying_i relations, we might conclude that the proposition Alice is saying_i at t, is the proposition that everything she is saying_j at t is untrue for some j distinct from i.

What about our original question? When we ask questions of the form 'what did so-and-so say?' are we asking about what has been said_1, or said_2, or what?[42] More importantly, how should we make sense of ordinary talk involving the word 'says' without the explicit subscripts, when semantic pluralism is being assumed? Here I think it is incredibly natural to think that the word 'says' is context sensitive and will pick out slightly different relations in different contexts.[43] The context sensitivity of semantic vocabulary was a tentative motivation for accepting semantic pluralism, but the converse also looks fairly attractive: once you've accepted semantic pluralism it's very natural to think that our ordinary semantic terms and terms for reporting speech acts will be context sensitive.[44] Putting this together, when you ask 'what is Alice saying?' the question you have asked depends on the context. If you are in a context where you have asked what Alice is saying_i then the answer can be the proposition that everything she is saying_j is false, provided j is distinct from i.

We can predict these results by making assumptions about the meanings of the subsentential components of Alice's utterance. For example, given the assumption that 'says' means_i the saying_j in Alice's context, and some simplifying disquotational assumptions (that 'Alice' means_i Alice etc) we can prove, via the compositionality principle mentioned earlier, that Alice's utterance of 'everything Alice is saying at t is untrue' says_i that everything Alice is saying_j at t is untrue.

Note, moreover, that the aforementioned compositionality assumption allows us to refute the following natural conjecture:

> 'says' means_i the saying_i relation.

For if it did then, given the same background assumptions, it would follow that at t Alice said_i that everything Alice said_i at t is untrue. But we rejected the latter hypothesis on the basis of Prior's theorem.

4.3 Other versions of Prior's paradox

Let me focus, for a moment, on one particular instance of Prior's theorem that seems particularly puzzling: the instance you get by substituting 'S' in Prior's theorem with 'it sounds as though Alice said at t that' ('Alice intended to say that' and 'Alice represents herself as knowing that' are similarly particularly puzzling instances).

---

false, but a slightly different proposition: that everything I'm thinking* is false.

[42] Presumably just as you stand in many different assertion-like relations to different propositions when you make a declarative utterance, you also stand in many inquisitive-like relations to different questions when you make an interrogative utterance. It's thus consistent to say that you've asked_1 what Alice said_1, asked_2 what she said_2 and so on. The conclusions I make below can still be drawn if we adopted a pluralism about asking as well as saying.

[43] Of course, even within a context it might be borderline which relation 'says' picks out, however I take this to be true of all context sensitive language.

[44] Aside: a sentence is context sensitive if different utterances of that sentence can be used to say different things. 'Context sensitive' in this sense is thus also context sensitive: in different contexts it can be taken to mean that different utterances of a sentence can be used to say_1 different things, to say_2 different things, and so on. Usually when a sentence is context sensitive in any of these senses it is context sensitive in all or most of them; the context sensitivity of the word 'says' is presumably not unusual in this regard and can be used in sentences to say_1 different things, to say_2 different things, and so on.

> If it sounds, at t, as though Alice said that everything it sounds as though she said at t is untrue, then there's something true and something false it sounds as though she said at t.

This instance of Prior's theorem is, I think, much more surprising and somewhat harder to solve than the original. For suppose that at t Alice utters nothing but the sentence 'everything at t it sounded as though I said is untrue'. Perhaps, as the no-proposition theorist would maintain, Alice does not succeed in saying anything by uttering this sentence. Be that as it may, it surely *sounds as though* she said something. It sound as though she said that everything it sounds as though she said is untrue. The straightforward analogue of the no-proposition response is completely implausible.

Note, however, that the paradox is puzzling for everybody. Since the antecedent of this instance of Prior's theorem is apparently satisfied, it follows that there are at least two things which it sounds as though she said at t. This also seems absurd – what is the other thing it sounded as though she said? We only heard her utter the words 'everything it sounds as though I said at t is untrue'.

Thus we must either conclude that it didn't sound as though Alice said that everything it sounded as though she said is untrue, or we must conclude there are two things it sounded as though she said. Absurd as this dilemma may sound it is just another deliverance of classical propositional logic and the logic of quantification into sentence position.

A good solution to the paradoxes should be also judged by its ability to deal with this variant paradox. It is clear, I hope, that whatever intuitive appeal the no-proposition view has with regard to the first version of Prior's paradox the advantages are completely lacking when it comes to this paradox.

The context sensitivity of 'says' predicted by the present view may help shed some light on this recalcitrant variant of Prior's paradox. To warm up to that, let us consider a simpler example: let us suppose that Bob and Caroline want to lift a box off a high shelf. Meanwhile David, in the next room, is telling his colleagues about how much his four year old, Enid, has grown. Caroline and Bob overhear David uttering the words 'Enid is tall'. Bob, not knowing that that David is talking about his child and wondering whether David is addressing them and suggesting that Enid could help them with their box, asks Caroline whether David just said that Enid is tall. Caroline responds negatively, and elaborates by telling Bob that David said that Enid is tall for a four year old, not that she was tall. Even though David uttered the words 'Enid is tall', this didn't result in him saying that Enid was tall since David was in a context in whith adult humans are not the salient comparison class.

I think a good case could also be made that it doesn't even sound to Caroline as though David said that Fred is tall. Given what we know about the context the utterance was made in, it sounds as though he was saying that Enid is tall-for-a-four-year-old – it's just obvious in this case that he wasn't saying that Enid is tall-for-an-adult, which is what 'Enid is tall' means relative to an ordinary context, or the context in which Bob and Caroline are talking, since it's clear to Caroline that Enid isn't tall in that sense (she cannot, for example, help them with their box). Presumably what David said and what it sounded as though he said to Caroline are the same in this case: that Enid is tall for a child, or something similar. Other scenarios are possible too: if we don't know enough about what context David is in, as Bob initially didn't, we won't know what he's said. In such cases there probably isn't a particular proposition you think he said, or a particular proposition which it sounds to you as though he said.

Of course, people can be bad at articulating how things sound to them. If you hear David utter the sentence 'Enid is tall' and you know it is a context different from your own, then you are typically in a position to know that he hasn't said that Enid is tall-for-your-context. However it is still very tempting to simply report David disquotationally, by saying 'David said that Enid is tall' or even 'it sounds to me as though David said that Enid is tall'. This might be tempting because the proposition David in fact said presumably cannot be described simply in your context because the word 'tall' doesn't express the appropriate property in your mouth. The proposition that Enid is tall, on the other hand, seems natural and simple because it's the proposition that I'm able to refer to in my context with a particularly simple definite description ('the proposition that Enid is tall'). Symmetrically from David's context that description would pick out his proposition, and it would be the proposition *I*

picked out with it that would be hard for him to articulate. If David were to try to describe the proposition that Enid is tall in his context he would also find he didn't have simple words for it – the words 'the proposition that Enid is tall' would not succeed unless what 'tall' means in his context is the same as what it means in mine.

Suppose that Alice utters, at t, the sentence 'everything it sounds as though Alice is saying at t is untrue' and this results in her saying some proposition, P, which, paralleling the above example, is distinct from the proposition that everything it sounds at though she's saying at t is untrue. Then, I maintain, not only did she say that P, but given that we are relevantly knowledgeable of the context in which she made her utterance, it sounded as though she said that P. In general we get to affirm the commonsense principle:

> In most circumstances what people have said and what it sounds as though they've said are the same.

If what people said and what it sounded as though they said frequently came apart then we would be constantly be misled about what people were saying. This would presumably prevent successful communication.

At any rate, this is enough to diffuse the paradox presented by Prior's theorem, for either we know enough about Alice's context to know that she's said that P (and not the proposition that everything it sounds at though she's saying at t is untrue.) In which case, I claim, P is what it sounds as though she's said, and not the proposition that everything it sounds as though Alice said at t is untrue. Or perhaps we don't know enough about the context to know what Alice has said, then I claim there's nothing it sounds as though she's said at t. Or perhaps in addition to sounding as though she said that P it also sounds as though she said that everything it sounds as though she said at t is untrue. All of these options are consistent with Prior's theorem.

Of course, if you ask someone what it sounded as though Alice said, they might well reply by uttering the sentence 'it sounded as though Alice was saying that everything it sounded as though saying at t was untrue (and nothing else)'. This kind of speech is not in general a good indicator of what things sounded like to that person. Analogously, if asked what it sounded as though Bob said, by uttering 'it sounded as though Bob said that Fred is tall'. Their mistake, I think, is not that they are bad at introspecting how things sound to them.[45] It is rather that when they introspect, and come to a correct belief about what things sounded like, they choose the wrong words to report this belief. As we have noted already, if Alice said that P, and even if the audience knows that Alice said that P, the audience might not be able to easily articulate this knowledge if they are in a different context from Alice; in such cases they may simply resort to a (false) disquotational principle in an attempt to articulate the thought.

4.4 Semantic Pluralism entails Radical Anti-Disquotationalism

While I have so far been using semantic pluralism to motivate failures of moderate disquotation, here I shall turn to showing that semantic pluralism entails far more radical violations of disquotational principles.

> Radical Anti-Disquotationalism: Hardly any of the cluster of important language-world relations relate utterances of 'snow is white' to the proposition that snow is white.

Of course, if 'snow is white' is context sensitive the idea that an utterance of 'snow is white' in context c says that snow is white is only plausible if c is a context relevantly like the context we are in right now. However my case for radical anti-disquotationalism does not rest on the context sensitivity of the sentence 'snow is white', for it will show that for any utterance of 'snow is white' (even ones made in this very context), most of the semantic relations will not relate that utterance to the proposition that snow is white.

[45] Although note that it's not entirely obvious that these kinds of facts are always introspectible; see Williamson [26].

The interest of radical anti-disquotationalism for our purposes is that it demonstrates that there is nothing exceptional about 'paradoxical' utterances and moreover it deflates a number of potential revenge arguments resting on weakenings of the disquotational principles.

The argument for this conclusion is actually surprisingly obvious. Suppose that says_1 and says_2 are two semantic relations. Suppose, moreover, that Alice makes an utterance of 'snow is white' and that the two semantic relations deliver differing verdicts for Alice's utterance: says_1 relates Alice's utterance to to the proposition that P and says_2 relations relates her to a distinct proposition, Q (where, lets say, P and Q differ over the minimum fluffiness that something needs to be to be snow). Since at most one of P and Q can be identical to the proposition that snow is white, at most one of the two semantic relations will relate Alice to the proposition that snow is white. The thought generalises, for if disquotational principles fail roughly half the time when there are only two relations, when there are several relations which do not agree with one another about Alice's utterance at most one of them can relate Alice to the proposition that snow is white, and thus most of them will not relate Alice to the proposition that snow is white.

Parallel arguments establish similar conclusions about reference.[46] There are plenty of mountain-like objects in the vicinity of Mt. Kilimanjaro that differ from one another only slightly over their boundaries – for example, some might include a small piece of dust near the bottom and others might not. Mt. Kilimanjaro is, of course, one of these objects, but there is nothing particularly special about it – it does not stand out as haloed among the other similar objects. There is presumably vagueness concerning the referent of 'Mt. Kilimanjaro'. It is natural to think there is also semantic multiplicity here – lots of extremely similar reference-like relations each relating the name 'Kilimanjaro' to one of these different mountain-like objects. I take it that none of these mountain-like objects has the unique privilege of being the most suitable semantic value of the name 'Kilimanjaro', not even Mt. Kilimanjaro, and that none of the variant reference relations plays a more important semantic role than any other. Yet since only one of these mountain-like objects is Mt. Kilimanjaro most of these reference relations will not relate the name 'Kilimanjaro' to Kilimanjaro.

It is initially quite tempting to turn this argument on its head and take it as a refutation of semantic pluralism. 'Look', one might argue, 'perhaps says_1 and says_2 play very similar roles, however if one of them is relating 'snow is white' to the proposition that snow is white and the other is not then surely the relation that does is the more important language-world relation, and we should not be concerned with the other relation in our semantic theorizing'. This argument can be repeated with each sentence of the language to show that if there are multiple important semantic relations, they must all agree with one another about which utterances relate to which propositions (i.e. they are extensionally equivalent). This, I take it, is enough to rob semantic pluralism of its interest.

Such arguments are liable to decieve. Although semantic multiplicity is not the same as context sensitivity, there is an analogous argument about contexts which is clearly flawed. The argument goes as follows: 'The word 'me' can be can be used in different contexts to pick out different people. In some contexts it picks out Alice, in some Bob and in others me. However there is clearly something theoretically distinctive about the last kind of context which makes it more deserving of the attention of semantic theorizing, for it is the only context in which 'me' behaves disquotationally.' The reason this argument is flawed, I take it, is that Alice or Bob could have verbally made the exact same argument and arrived at very different conclusions.

Something similar is going on in the case of semantic multiplicity. The problem in the above argument was that I used (as opposed to mentioned) a word, 'me', that had several equally good referents associated with it; in the present case we used an expression, 'the proposition that snow is white' that while not context sensitive, is associated with many different propositions including P and Q. Now suppose I make the following announcement:

> Relations that relate utterances of 'snow is white' to the proposition that snow is white are better than those that don't.

[46] Here I draw on the discussion in Hawthorne [14].

Since 'snow is white' says$_1$ that P and says$_2$ that Q, then we can ask what I've said$_1$ and said$_2$ when I make the above announcement. Assuming some compositionality, I've said$_1$ that relations that relate utterances of 'snow is white' to P are better than those that don't and I've said$_2$ that relations that relate utterances of 'snow is white' to Q are better than those that don't.[47] Not only is there a perfect symmetry here between P and Q, but I've said$_1$ and said$_2$ contradictory things, which explains why it might be bad to make announcements like this.[48]

Here is another way of seeing this point. Let us suppose that either 'snow is white' says$_1$ that snow is white or that it says$_2$ that snow is white (but not both.) Which saying relation is the distinguished one depends which of these two disjuncts is true. However there are two notions of truth that we can apply to an utterance: an utterance is true$_1$ if it says$_1$ a truth and no falsehoods ($\exists P(S_1 uP \wedge P) \wedge \forall P(S_1 uP \rightarrow P)$) and true$_2$ if it says$_2$ a truth and no falsehoods ($\exists P(S_2 uP \wedge P) \wedge \forall P(S_2 uP \rightarrow P)$) – for all we've said the first disjunct is true$_1$ and false$_2$ and the second disjunct is true$_2$ and false$_1$ (or perhaps neither.)

Thus the thought that there is a unique proposition that bears a distinguished relation to the words 'snow is white' is false. Neither P nor Q bears a more distinguished relation to this sentence than the other. Thus the logically stronger claim that the proposition that snow is white uniquely bears this special relationship to the words 'snow is white' is also false.[49] P's claim to being special is just as good as Q's – this parity remains even if P is the proposition that snow is white while Q isn't (or vice versa.) Again, we should not get too worked up about this for 'P is the proposition that snow is white' might say$_1$ a truth and say$_2$ a falsehood and 'Q is the proposition that snow is white' might say$_2$ a truth and say$_1$ a falsehood.

Although I have stressed that semantic multiplicity is not the same as context sensitivity, the analogy with context sensitivity is not entirely misleading here. Let us suppose that 'snow is white' is not context sensitive, and that there's a unique proposition said by all utterances of this sentence. Even so, the description 'the unique proposition said by all utterances of 'snow is white'' is context sensitive due to the context sensitivity of 'says'. In some contexts it denotes, P, the unique proposition said$_1$ by all utterances of 'snow is white' in others it denotes Q, the unique proposition said$_2$ by all utterances of 'snow is white'.

The key to making ones peace with these anti-disquotational conclusions is to recognise that there is nothing special about Mt. Kilimanjaro over the other mountain-like entities in its vicinity, and that the proposition that snow is white is no more important than the plethora of propositions distinct from it but with very similar modal profiles.

Lastly, let me stress that RAD still leaves room for the acceptability of some disquotational *utterances*. For example, although most semantic relations do not relate 'snow is white' to the proposition that snow is white, it could still be true that some or all of these semantic relations relate utterances of the the *disquotational sentence* "snow is white' means that snow is white' to a true proposition. In which case, even though RAD is true, it could turn out that certain utterances stating disquotational principles will come out true in some or all the different senses of true. For example, a plausible sufficient condition for an utterance of this sentence to be true$_i$ would be if 'snow is white' meant$_i$ and meant$_j$ the same proposition, where meaning$_j$ is the relation 'means' means$_i$. More explicitly:

1. 'means' means$_i$ meaning$_j$ (this holds for some j according to the present view)

2. 'snow is white' means$_i$ and means$_j$ the same proposition: P.

[47] Actually I've made quite a few assumptions here to ease readability. I've assumed, for example, that 'utterances' refer$_1$ and refer$_2$ to utterances, when it could be referring$_1$ and referring$_2$ to noises that are only slightly different from utterances. Similar assumptions were made about the other words used in my announcement.

[48] For all I've said it's possible in some circumstances to make an utterance and say$_1$ something without saying$_2$ anything. Even if this is what is happening here, it seems somewhat arbitrary to think that relations that relate 'snow is white' to P are better than those that don't.

[49] This fallacy can be disguised in lots of different ways. For example, it might seem as though the set of worlds at which snow is white is the uniquely most natural set of words to associate with the sentence 'snow is white'. However there are variant understandings of 'world at which' which generate other sets of worlds which bears an equally natural relationship to this sentence.

3. "snow is white' means that snow is white' means$_i$ that 'snow is white' means$_j$ that P.

3 follows from 1 and 2 and compositionality given the simplifying assumption that the quotation name ''snow is white'' means$_i$ the sentence 'snow is white', and similar such assumptions. Given 2 'snow is white' means$_j$ that P and so given 3, "snow is white' means that snow is white' means$_i$ a truth, namely that 'snow is white' means$_j$ that P.

Liar-like paradoxes will certainly place limitations on the extent to which even this form of disquotation can hold.[50] However it's entirely possible that disquotational utterances of sentences that do not involve semantic vocabulary are acceptable. The main point to bear in mind is that the justification for why these utterances are acceptable, as we saw above, is not at all routine, and so we have at least a clearer picture of what happens when disquotational utterances involving paradoxical sentences fail to be true$_i$. In these cases, the meaning$_i$ and meaning$_j$ relations disagree about the proposition assigned to the paradoxical utterance (here meaning$_j$ is what 'means' means$_i$).

4.5 Revenge Paradoxes

Let us now take a moment to briefly recap the main features of the view I have proposed, with an eye to the prospects of generating a revenge paradox for it.

Recall that according to our definition of utterance truth in section 1, an utterance is true if it says something true (formally $\exists P(u \text{ says that } P \wedge P)$). Let u be an utterance of the sentence 'u is not true'. Now by appealing to the principle Utterance Truth (which can now be proven from our definition of utterance truth and the assumption that u says at most one thing) we can infer that u does not say that u is not true.

On the view I have proposed, u says something else instead. As we have seen, semantic pluralism helps explain this phenomenon. Firstly, it removes the air of mystery surrounding the question of which proposition u *does* in fact say, if not that u is not true. The saying relation isn't the only important language-world relation, and consequently truth so defined isn't the only semantically important property of utterances. u might say that u is not true*, where truth* is defined in terms of another one of these semantic relations, saying* (i.e. truth* is defined as $\exists P(u \text{ says}^* \text{ that } P \wedge P)$).

Secondly it predicts, on quite general and principled grounds, failures of the sort of disquotational reasoning that gets us into trouble with the liar. For if there are many semantic relations relating utterances of 'u is not true' to lots of slightly different propositions then at least some, and maybe even most, will not relate utterances of 'u is not true' to the proposition that u is not true (they can't relate an utterance of 'u is not true' to different propositions and at the same time be such that they all relate that utterance to the proposition that u is not true). It should thus be of no surprise that a particular one of these – the *saying* relation – does not relate u to this proposition; indeed the term 'the saying relation' is, on one plausible theory, a context sensitive term for a particular one of these relations.

However this does raise a natural question: do *any* of the relevant semantic relations relate the utterance u to the proposition that u is not true. (After all, we maintained that some of these semantically significant relations related utterances of 'snow is white' to the proposition that snow is white, even if not all of them do.) And do any revenge paradoxes ensue if we assume that there are semantic relations relating 'u is not true' to the proposition that u is not true?

Note that it is perfectly consistent to say that some semantic relation, says*, relates u to the proposition that u is not true. The most we can deduce *a priori*, from limitative results such as Prior's theorem, is that u does not *say* that u is not true (and, by parity, the most we can deduce *a priori* about what u says* is that it doesn't say* that u is not true*).

However, although no paradox ensues if we assume the 'disquotated proposition' (so to speak) is among the propositions said in these particular cases, we might attempt to generalize this idea. We

[50] For example, there are some instances of "S' is true if and only if S' that are classically inconsistent, and so assuming that the expressing relations are reasonably well-behaved no expressing relation can relate utterances of this sentence to truths.

might attempt to justify a much more general and far-reaching disquotational schema that applies to *any* sentence of our language:

> VERY MODERATE DISQUOTATION: for any utterance of 'ϕ' there is some semantic relation that relates that utterance to the proposition that ϕ.

Although I have expressed sympathy to a few instances of this schema (when 'ϕ' is 'snow is white', for example), the full schema does not follow from anything I have said so far: to accept it would be an additional substantive commitment. Indeed, once we have properly recognized the pervasiveness of failures of disquotational principles, it becomes hard to motivate on general grounds. However, let us for the sake of argument see what would happen if we did accept it.

Say that a proposition P is said$^+$ by an utterance u if some relation playing the expressing-role relates u to P. Given this definition, VERY MODERATE DISQUOTATION amounts to the schema: utterances of 'ϕ' say$^+$ that ϕ. Now let L^+ be the sentence 'no utterance of L^+ says$^+$ a truth and no falsehood' 'no utterance of L^+ says$^+$ a truth and no falsehood' (this is clearly just a variant of the paradoxical sentence L considered in section 3.2). Substituting this sentence into our disquotational principle, and appealing to the identity L^+ = 'no utterance of L^+ says$^+$ a truth and no falsehood', we get:

(*) Utterances of L^+ say$^+$ that no utterance of L^+ says$^+$ a truth and no falsehood.

But note that by exactly the same reasoning we appealed to in section 3.2 (spelled out formally in the appendix) the above entails that no utterance of L^+ says$^+$ a truth and no falsehood.

If we assume, in line with the proposed solution, that utterances of L^+ say something then it follows that an utterance of L^+ says$^+$ something, and thus that they say$^+$ at at least one falsehood (because we have just proved that no utterance of L^+ says$^+$ a truth without also saying$^+$ a falsehood, and so no utterance of L^+ can say$^+$ anything without saying$^+$ a falsehood).

How bad is this conclusion? Notice firstly that it is hardly surprising to find out that some utterances say$^+$ many different things, for it is the main contention of the semantic pluralist that most utterances bear lots of related expression-like relations to lots of different propositions. But the fact that utterances of L^+ say$^+$ falsehoods is problematic for the very kinds of reasons we discussed in section 3.2. Anyone accepting (*) must also accept the conclusion we proved from it: that no utterance of L^+ says$^+$ a truth and no falsehood, and the most natural way to communicate this is to produce an utterance of L^+.

Suppose that saying$_i$ is a relation by which utterances of L^+ say$_i$ a falsehood. Then, since we derived L^+ from (*) and some background assumptions that plausibly say$_i$ truths, it ought to follow that the proposition (*) says$_i$ is false (here we are appealing to the idea that saying$_i$ a truth is preserved by logical consequence[51]). Now the saying$_i$ relation plays the expressing role, and presumably part of that role is that one shouldn't say$_i$ (or assert$_i$) things that one doesn't know. Since one cannot know a falsehood, we shouldn't say$_i$ it, and since by assumption uttering (*) in an assertive tone of voice results in one assertively saying$_i$ it, we shouldn't utter (*) in an assertive tone of voice.[52]

Thus, I claim, we should not assertively utter (*), which is a particular instance of VERY MODERATE DISQUOTATION governing the sentence L^+. It is natural to ask for a deeper explanation for why this particular instance cannot be uttered assertively, especially when other instances, including those concerning the ordinary liar, appear to be unproblematic. We might obtain a little more clarity if we explicitly calculated what instances of VERY MODERATE DISQUOTATION say$_i$, so that we can see explicitly what it would take for it to be false.

Begin by noting that the notion of an 'expressing-like' relation was defined in vague and context sensitive terms; it was introduced as a relation that satisfied the expressing-role for a given language,

[51] This idea can be justified more rigorously by appealing to the compositionality of the saying$_i$ relation.

[52] According to an alternative version, attitudes verbs like 'knows' are also the subject of semantic pluralism, so there are lots of knowledge-like attitudes hanging around (see our remarks on attitudes in section 4.1). However it's natural to think that in this setting the expressing role should be that one shouldn't assert$_i$ what one doesn't know$_i$; by assuming the factivity of knowledge$_i$ the above reasoning goes through exactly the same.

a role which was itself vaguely specified. Thus it is exceedingly natural to think that the term 'expressing-like' is just as much subject to semantic pluralism as the word 'saying': there are lots of different clouds of similar relations that could be the denotation$_j$ (for some j) of the expression 'expressing-like relation'.

Although the following considerations will not turn heavily on the details, here is one natural picture of how these relations might be structured. For each saying relation, says$_j$, there is a cloud of closely related relations that play the roughly same role in communication as saying$_j$ does: call the set of indices that index such relations $S(j)$. Any cloud of closely related relations generated by some saying relation in this way is a natural candidate to be one of the clouds of relations corresponding to the term 'expressing-like relation'. Indeed, it is extremely natural to go so far as to think that if 'says' denotes$_i$ the saying$_j$ relation, then 'expressing-like relation' denotes$_i$ the set of relations indexed by $S(j)$ (i.e. the relations playing roughly the same role as says$_j$). (This generalizes the thought that we might express ambiguously with 'an expressing-like relation is one that plays roughly the same role in communication as the saying relation'.)

Another natural thought is that whenever i and j are distinct indices, since the saying$_i$ and saying$_j$ relations are slightly different, the clouds of relations they generate, and thus $S(i)$ and $S(j)$, will be also be slightly different. Now recall that by Prior's theorem we argued that the word 'says' does not denote$_i$ the saying$_i$ relation, it denotes a slightly different relation, saying$_j$ (belonging to the i-cloud of relations, $S(i)$). So by the above natural thought we can analogously conclude that the term 'expressing-like relation' applies$_i$ to those relations that are sufficiently similar to says$_j$.

We are now in a position to calculate what a given instance of Very Moderate Disquotation would say$_i$. Suppose that 'ϕ' is a sentence which says$_i$ the proposition that P, and 'expressing-like relation' denotes$_i$ a property that characterizes the set of relations indexed by $S(j)$. Assuming (for convenience) that 'any' means$_i$ any, 'utterance' means$_i$ utterance, and so on, it follows by compositionality that the ϕ instance of Very Moderate Disquotation says$_i$ that:

> For any utterance of 'ϕ' there is some expressing-like$_j$ relation that relates that utterance to the proposition that P.

Let us now consider whether an instance of Very Moderate Disquotation could say$_i$ a falsehood. In other words, could the above claim fail? That would mean that there is no expressing$_j$-like relation that relates utterances of 'ϕ' to P.

Note in particular that since saying$_i$ relates 'ϕ' to P it follows that says$_i$ cannot be an expressing$_j$-like relation. A more compact terminology can be introduced: let $\sigma(i)$ represent the index of the saying relation that 'says' denotes$_i$. (So, for example, if 'says' denotes$_i$ the saying$_j$ relation then $\sigma(i) = j$.) Then in this terminology we have a counterexample to the principle that $i \in S(\sigma(i))$.

By contrast, we argued earlier that the word 'says' denotes$_i$ not the saying$_i$ relation but sufficiently similar relation, saying$_{\sigma(i)}$ which in our more compact notation says: $\sigma(i) \in S(i)$. This shows that the relation that holds between i and j when the index i satisfies the expressing$_j$-role is not a symmetric relation: that is we have $j \in S(i)$ but not $i \in S(j)$, where $j = \sigma(i)$.

This could happen in a number of different ways. On the model we have been considering satisfaction of the expressing$_i$ role amounts to being a relation that is sufficiently similar (in some to-be-specified sense) to the saying$_i$ relation. Asymmetry could arise if what counts as 'sufficiently similar' to says$_j$ (i.e. what it takes to satisfy the expressing$_j$-role) is more stringent than what counts as sufficiently similar to says$_i$ (i.e. what it takes to satisfy the expressing$_i$-role).

In summary, then, the most obvious revenge paradox does not arise for the present view: in particular, the view we have defended is not committed to Very Moderate Disquotation. We have seen that instances of this schema can bear semantic relations to falsehoods when there is asymmetry in the relation *i satisfies the expressing$_j$-role*. Since it is not implausible that this relation is asymmetric, I think we have no good motivation to accept even Very Moderate Disquotation; no revenge paradox that appeals to it as a premise should therefore trouble us.

Let me end our discussion by briefly remarking upon the scope of our approach to the paradoxes. Semantic pluralism clearly has some kind of diagnosis to offer for the properly *semantic* paradoxes –

paradoxes involving meaning, saying, truth, reference, and so on, and which can also can be construed broadly to include the paradoxes of thinking and other propositional attitudes. However semantic pluralism is clearly not going to illuminate a strand of paradoxes which, following Frank Ramsey, we might call the *logical paradoxes*. By contrast with the semantic paradoxes, which begin with some kind of premise about a semantic notion, the logical paradoxes are formulated in purely logical terms.[53] In these cases logic tends to decide the answers for us.

If, for example, there were a proposition, p, identical to the proposition that p is not true, then we would be in a familiar kind of trouble, and since neither language-world nor mind-world relations play a role in this paradox semantic pluralism has nothing to offer by way of diagnosis.

There are certain theories of propositions — specifically theories of structured propositions — which are inconsistent because they prove versions of the diagonal lemma, and so allow one to construct self-referential propositions.[54] But it should be noted that the way of talking about propositions we have adopted here, formalised in a fragment of higher-order logic using quantification into sentence position, is not inconsistent. As we noted earlier a propositional truth predicate is not needed in this setting. One can, however, introduce a truth *operator*. It is a truth functional operator (whose truth table just takes T to T and F to F) so it follows that an operator with this truth functional behaviour can already be defined from the standard truth functional connectives: we could, for instance, define Tp as $\neg\neg p$.[55] In this setting the aforementioned paradox involving propositions becomes a theorem to the effect that no proposition is identical to the proposition that it is not true.

THEOREM: $\neg\exists P(P = \neg TP)$

This theorem is proved by noting that Leibniz's law entails that $P = \neg TP$ implies $P \leftrightarrow \neg TP$, and then noting that this is inconsistent with $TP \leftrightarrow P$ (this last biconditional is guaranteed by our definition of T).

Although higher-order logic on its own is consistent, it is inconsistent if supplemented with principles that encode the idea the propositions are structured.[56] So much the worse for structured theories of propositions. Alternative doctrines about fineness of grain such as *Booleanism* — the view that Boolean equivalents are identical — are straightforwardly consistent in higher-order logic. (Booleanism is validated in the models we consider in the appendix.)

Although this way of capturing proposition talk in higher-order logic is consistent, one might wonder about the status of theories that treat propositions as individuals and invokes a truth predicate. A general framework (that is neutral about Booleanism[57]) can be formulated in the higher-order setting by introducing a functor, pronounced 'that', which takes a sentence as argument and produces a singular term naming a proposition as a result, and a predicate 'true' which takes the name of a proposition and gives a sentence back. (In this case we say that *that* has type $t \to e$, and *true* has type $e \to t$, where t is the type of a sentence and e the type of a singular term.) A theory of first-order propositions can thus be stated in a language containing *that* and *true*, and axiomatised with an axiom

[53] Ramsey understood 'logic' broadly enough to include set theory, and thus counted the set theoretic paradoxes as logical. I'll be focusing on puzzles that are purely logical even by todays standards.

[54] See Whittle [24]. There is a close relation between the proof that the diagonal lemma can be reconstructed in a structured proposition theory and the Russell-Myhill paradox which also shows the inconsistency of certain theories of structured propositions. For any structured proposition r which is of the form (O, p) for some operator O, one can define a proposition $Sub(p, q, r)$ which is true iff the result of substituting q for p in r is true: $Sub(p, q, r) := \forall X(r = Xp \to Xq)$ where X ranges over operators). Given this definition of substitution one can reconstruct a form of the the diagonal lemma, and one gets a liar proposition if one lets O be negation. The resulting liar proposition is moreover the same, modulo a little logical manipulation, as the proposition Myhill constructs to show the inconsistency of structured theories of propositions. For a recent discussions of the Russell-Myhill paradox see Hodes [REF] section [REF] and Dorr [REF] section [REF].

[55] In the setting of higher-order logic with λ-expressions, one may simply define a truth operator as the identity operator, $\lambda p\, p$, which maps each proposition to itself. (Note that in a structured setting identity and double negation are not the same operator since they have different structure. However, they do have the same truth-functional behaviour.)

[56] Specifically, the principle Dorr [6] calls *Structure* is inconsistent in higher-order logic: $\forall XY\forall xy(Xx = Yy \to (x = y \wedge X = Y))$ where X and Y range over operators (type $t \to t$) and x and y over propositions (type t).

[57] The theory is, of course, not neutral about the naïve version of the structured proposition theory: it is inconsistent with it, but only because that theory is already inconsistent.

saying that *true* is a left inverse of *that*: that is when *true* is composed with *that* you get the identity function (of type $t \to t$). In higher-order logic can be concisely captured with the axiom:

$$\lambda P\, \text{true}(\text{that}(P)) = \lambda P\, P).$$

Given the assumption that the *that* functor is injective (an assumption one can state precisely in higher-order logic) one can show that *that* has a left inverse, and so we can always construct such a truth predicate given a choice for the *that* functor.[58] In this theory *true* plays the role of a truth predicate, and one can prove a version of the the T-schema: $true(that(P)) \leftrightarrow P$ (the right-hand side is simply identical to the left-hand side). One can likewise prove that there are no liar propositions: $\neg\exists x(x = that(\neg true(x)))$.

Returning to propositionally quantified logic, note that the result that there are no self-referential propositions was proved by logic alone. There is no non-logical premise concerning, for example, linguistic truth or expressing, and so no diagnosis is needed beyond the proof of this fact.[59] So regarding the logical paradoxes of this sort we are *all*, in some sense, 'no-proposition theorists' (unless we are logically revisionary). In this case, at least, making such a move is akin to denying the existence of a largest number, or to denying the existence of a set that contains itself if and only if it doesn't. But most importantly, such responses to the logical paradoxes do not seem to be subject to the same expressive problems that arise for the no-proposition account of the semantic paradoxes; and it is exactly here that semantic pluralism offers a preferable alternative.

5 Appendix: A Model of Semantic Pluralism

Let me summarize the main features of the view sketched in the last section:

1. There are multiple language-world relations that play roughly similar roles in explaining how language is used to communicated beliefs.

2. All of these relations are compositional, and relate each utterance to a unique proposition (with the usual exceptions for failed demonstratives, and similar phenomena).

3. Few (if any) of these relations relate utterances of 'P' to the proposition that P in any context, and few relate utterances of "P' says that P' to a true proposition in many contexts.

4. We can, without fear of paradox, theorize completely generally about the different meanings of each expression, even expressions which are related to semantic relations by some of the semantic relations.

It is the last point that needs some further defence. For although I have argued that disquotational principles are not motivated on this view, and so the standard assumptions required to derive paradoxes are blocked, it is hardly a general argument that the resultant theory is paradox free. To do that we would firstly have to have a more explicit theory than the merely suggestive remarks made in the last section, and secondly we would have to have a consistency proof of that theory. This is what I shall try to do here.

The idea will be to formulate a language containing expressions corresponding to the indexed 'expressing$_i$' relations we have been informally employing here. These indexed expressions don't correspond in an obvious way to colloquial English, however we do know that they are connected to uses of the word 'says' in certain contexts. A natural way to understand the indices, then, is to think

[58] Moreover, this theory is easily seen to be consistent. Its models all have the feature that there are at least as many things in the domain of the first-order quantifiers, D^e, as in the propositional quantifiers, D^t. In these models *that* is interpreted as an injection from D^t to D^e, and *true* is interpreted as a left inverse of the interpretation of *that* from D^e to D^t.

[59] Of course, sometimes people ask for diagnoses of surprising theorems (such as, say, the Banach-Tarski paradox). But these are usually requests for more intuitive proofs, or for more intuitive explanations of the terms in the theorem.

of them as anaphoric on particular uses of the colloquial word 'says'. An analogy is perhaps helpful here: if at time t I had used the word 'tall' to mean tall relative to some comparison class X, and I wanted to refer back to it frequently without having to precisely describe the relevant comparison class I could introduce a context insensitive expression 'tall$_t$' stipulated to be used the same way that 'tall' was at t.

The language must also permit quantification into sentence position and some means of stating the identity of two propositions. The language should also be capable of talking about, or at least representing, its own syntax – it is fairly standard to use the language of arithmetic to achieve this. In this language we can represent indices, sentences, and any other theoretically important entity, using numbers. Let our language, $\mathcal{L}$, then be the language of first order arithmetic with the following further primitive vocabulary:

- A binary relation, $[\![S]\!]_i$. Once the singular terms S and i are supplied, $[\![S]\!]_i$ grammatically takes the place of a sentence.
- A binary modal connective $\equiv$ representing propositional identity.
- An infinite stock of propositional variables, and the propositional quantifier $\forall p$.
- An optional stock of non-logical n-ary predicates P^n_i.
- An optional stock of n-ary non-logical connectives, C^n_i.

Note that in order to talk about what a sentence says$_i$ we have adopted a binary relation, such that when you provide a sentence and an index, i, the resulting sentence has the same truth conditions as the proposition said$_i$ by that sentence: we write this $[\![S]\!]_i$. One may intuitively read $[\![S]\!]_i$ as 'what would be said$_i$ by S in context c in language $\mathcal{L}$'. However, note that this paraphrase is not quite accurate: in a world where 'grass is green' had been used to mean that snow is white, what would have been said by 'grass is green' in a context might have been not have the same truth value $[\![$ 'grass is green' $]\!]^c$ – if we are to paraphrase our primitive this way, must think of our description as picking out the proposition that would be said by an utterance of S as *rigidified* to the way we *actually* use the sentence S.

This formalism makes for a neater presentation than if we had gone for a three place primitive, 'utterances of S say$_i$ that P (in c)', that takes two singular terms and a sentence as arguments and results in a sentence. The advantage of the latter primitive is that the substantive assumption that every sentence says$_i$ a unique proposition in each context is not built straight into the formalism. However, since I have argued for this assumption already, it is natural to prefer my primitive.

Strictly speaking we would also want to represent context sensitivity by relativising $[\![\cdot]\!]_i$ to a context in addition to an index. There is a fairly natural way to do this within the model I develop shortly, however it distracts from the main insights and isn't necessary in our present language which only contains the context insensitive language of arithmetic and context insensitive expressions for the different expressing$_i$ relations. (Contexts would become more interesting if we wanted to add a context sensitive word corresponding to 'says', which is context sensitive$_j$, for each j, between the different saying$_i$ relations.)

A natural simplificatory assumption would be to suppose that necessarily equivalent propositions are identical and to define the $\equiv$ connective from a modal operator with the definition $\Box(A \leftrightarrow B)$. Nothing turns on this, it is purely for the sake of having a concrete model for the relation – one can also do everything that follows with a primitive propositional identity connective.

We shall also adopt some fairly standard conventions with respect to representing syntax in arithmetic. We assume a Gödel numbering, with the numeral for the Gödel number of a formula ϕ being written $\ulcorner\phi\urcorner$. There are also primitive recursive functions $\dot{\neg}, \dot{\wedge}, \dot{\vee}$ and so on taking the Gödel number of some formulae to the Gödel number of the result of combining negation, conjunction or disjunction to those formulae. We write $x[\dot{v}/y]$ to represent the Gödel number of the result of a substitutions in the formula x is a Gödel number of. In arithmetic we can define predicates $Sent$, At and Ver to

saying that a number is the Gödel number of a sentence, atomic arithmetical sentence and an atomic arithmetical truth respectively. Since the domain of discourse in our models will just be the set of natural numbers we have a straightforward way to represent and quantify over indices.

Let me now describe a formal system that allows us to derive theorems that capture some of the ideas we have mentioned above. The theorems of this system are themselves say$_i$ different things.

CL The theorems of classical quantificational logic and the axioms of propositional quantificational logic.

S5 The axiomatization of S5 governing the modal operator $\Box$. The necessitation rule can be applied to anything provable in T.

PA The axioms of Peano arithmetic with full induction (i.e. including instances of induction including non-arithmetical vocabulary.)

At. $\forall i \forall x(At(x) \rightarrow (\llbracket x \rrbracket_i \equiv Ver(x)))$

C$\rightarrow$. $\forall i \forall x \forall y(Sent(x) \wedge Sent(y) \rightarrow (\llbracket x \dot{\rightarrow} y \rrbracket_i \equiv (\llbracket x \rrbracket_i \rightarrow \llbracket y \rrbracket_i))$

C$\vee$. $\forall i \forall x \forall y(Sent(x) \wedge Sent(y) \rightarrow (\llbracket x \dot{\vee} y \rrbracket_i \equiv \llbracket x \rrbracket_i \vee \llbracket y \rrbracket_i))$

C$\wedge$. $\forall i \forall x \forall y(Sent(x) \wedge Sent(y) \rightarrow (\llbracket x \dot{\wedge} y \rrbracket_i \equiv \llbracket x \rrbracket_i \wedge \llbracket y \rrbracket_i))$

C$\forall x$. $\forall i \forall x(Sent(x(\bar{0}/v)) \rightarrow (\llbracket \dot{\forall} v x \rrbracket_i \equiv \forall y \llbracket x[\dot{y}/v] \rrbracket_i))$

C$\forall p$. $\forall i \forall x y(Sent(x(\ulcorner 0 = 0 \urcorner / q) \wedge Sent(y)) \rightarrow \Box(\llbracket \dot{\forall} p x \rrbracket_i \rightarrow \llbracket x[\dot{y}/q] \rrbracket_i))$

C$\neg$. $\forall i \forall x(Sent(x) \rightarrow (\llbracket \dot{\neg} x \rrbracket_i \equiv \neg \llbracket x \rrbracket_i))$

C$\Box$. $\forall i \forall x(Sent(x) \rightarrow (\llbracket \dot{\Box} x \rrbracket_i \equiv \Box \llbracket x \rrbracket_i))$

Note the we are assuming, roughly, that 'and' means and, 'not' means not and so on. These stipulations seem plausible because 'not', 'and' and so on do not seem to be subject to the same kinds of issues we raised for words like 'snow' and 'white'.[60]

Call the above theory T. Note that in assertively uttering the theorems of T I commit myself to their truth$_i$ for each i: part of the role that each truth predicate plays includes the role of regulating assertive utterances. We thus might also want to state that theorems of the above system are true$_i$ for all i. Note that this claim is not treated as an axiom of our system since it is not import for doing semantics: our theory clearly isn't out to document all truths – for example, it isn't supposed to decide whether there is life on mars – it is only supposed to be a simple system for doing some basic semantics. It is not at all clear that the claim that the theorems of T say truths in all contexts is important for doing semantics. Nonetheless, one would hope at least that the resulting set of sentences is consistent: I am not only asserting the theorems of T, I also asserted that the theorems of T are true$_i$ for all i.

Truth$_i$ can be defined from the notion of saying$_i$ 'For all p if an utterance of S says$_i$ that p in c, then p'. In our formalism, with the 'what is said' functor, the quantification is unecessary – the truth of what S says$_i$ can be stated just by asserting what is said: $\llbracket S \rrbracket_i$.[61] We can introduce the idea that the theorems of T are true$_i$ for all i with the following schema:

[60] Observe also that the clause for the propositional quantifier differs slightly from the clause for the singular quantifiers. This is because, in the present language, we can understand universal quantification over numbers substitutionally, whereas the substitutional interpretation is not necessarily appropriate for the propositional quantifiers. This issue could be solved by enriching our primitives to include a notion of semantic expression relativised to a kind of higher order variable assignment. This is not a complication that needs to be attended to for our present purposes.

[61] Note on one interpretation, asserting that 'snow is white' is true is different from simply asserting the semantic value of 'snow is white' (i.e. asserting that snow is white) because there can be worlds in which snow is white, even though 'snow is white' is not true because means something other than that snow is white. I think that here we should just think of the relevant notion of truth in a langueg L as being truth in L as we actually use it.

$\forall i[\![\ulcorner\phi\urcorner]\!]_i$, where ϕ is any theorem of T.

Call the resulting theory T_1. One might also want to assert that all the theorems of T_1 are true in all contexts, one can clearly iterate this any number of times. Let T_{n+1} be the theory you get by adding the following rule to T_n:

$\forall i[\![\ulcorner\phi\urcorner]\!]_i$ whenever ϕ is a theorem of T_n.

What is the status of the stronger theories T_n where n is large over, say, T_1 or T? If you make an utterance of a sentence, S, according to the expressing role, that utterance is permissible only if S is true$_i$ for each i. However, since there is in general no good reason to think that whenever something is assertable, it should be assertable that it's assertable, there is generally no reason to think that we should accept the theorems of T_n for arbitrarily large n.

5.1 Consistency proof

A model of our language is effectively a fixed domain Kripke model, augmented in a natural way to deal with propositional quantification. Necessity can be interpreted in the natural way using quantification over all worlds, propositional quantification by quantification over arbitrary sets of worlds, the semantic value function will determine a function from worlds to truth values for each sentence and index and the arithmetical vocabulary should recieve standard interpretations at each world. If non-logical connectives are included they can simply be modelled by a function that assigns to each world a set of sets of worlds.

We shall therefore want to construct a class of Kripke models. Each member of this class will agree about everything except the interpretations of $[\![\cdot]\!]_i$. Each model will therefore consist of:

- A fixed set of worlds, W.
- A domain D, is the set of natural numbers $\mathbb{N}$.
- The interpretation function, I, assigns, for each world, an extension to the non-logical vocabulary.
  - The interpretation of the arithmetical vocabulary at any given world is standard.
  - $I(P_i^n, w)$ is a subset of D^n – i.e. an n place relation over the naturals.

Finally, we need to supply the extension of $[\![\cdot]\!]$.. Technically, since it is a binary relation, its interpretation should be a function from worlds to sets of ordered pairs. However, there is a more intuitive way to think about it. Relative to an index, i, $[\![\cdot]\!]_i$ determines a function that takes a sentence (or, more accurately, the Gödel number of a sentence) and returns a set of worlds representing the semantic value$_i$ of that sentence. In other words, $[\![\cdot]\!]_i$ takes a sentence and outputs the proposition that would be said$_i$ by the sentence.[62] The possible interpretations of the saying function $[\![\cdot]\!]_i$ relative to a fixed index are given by the set of functions from Gödel numbers of sentences to propositions:

$S := \{f \mid f : \mathbb{N} \to \mathcal{P}(W)\}$.

Each element of S represents a semantic value function. To complete our interpretation we need to pick for each index, i, a member of S, to represent the interpretation of $[\![\cdot]\!]_i$. In our setting contexts are being represented by numbers so we just need an ω sequence of elements of S to fully interpret our language.

Sentences of the language, then, will be evaluated at triples of worlds, elements of S^ω, and variable assignments. A variable assignment is a function on the set of singular and sentential variables that

[62] A function f from sentences and indices to propositions is equivalent to a function from worlds to ordered pairs: just let $g(w) = \{\langle\ulcorner\phi\urcorner, i\rangle \mid w \in f(\phi, i)\}$.

maps singular variables to natural numbers and sentential variables to subsets of W. I will assume that assignments can be extended so as to assign values to complex singluar terms in the standard way. If v and u are two assignments write $u[x]v$ and $u[p]v$ to mean that u and v agree about every variable except, possibly, at x/p. I determines the truth value of an atomic sentence at a world in the standard way, and truth at world and assignment (written $w, v \Vdash \phi$) can be extended to complex formulae as follows:

- $w, s, v \Vdash \neg\phi$ iff $w, s, v \nVdash \phi$
- $w, s, v \Vdash \Box\phi$ iff $w', s, v \Vdash \phi$ for every $w' \in W$
- $w, s, v \Vdash \phi \wedge \psi$ iff $w, s, v \Vdash \phi$ and $w, s, v \Vdash \psi$.
- $w, s, v \Vdash \phi \vee \psi$ iff $w, s, v \Vdash \phi$ or $w, s, v \Vdash \psi$.
- $w, s, v \Vdash \forall x\phi$ iff $w, s, u \Vdash \phi$ for every $u[x]v$.
- $w, s, v \Vdash \forall p\phi$ iff $w, s, u \Vdash \phi$ for every $u[p]v$.
- $w, s, v \Vdash [\![t_0]\!]^{t_1}$ iff $w \in s(v(t_1))(v(t_0))$

In the last clause note that $s(v(y))$ is the semantic value function of the index corresponding to $v(y)$ and $s(v(y))(v(x))$ is the proposition the sentence whose Gödel number is $v(x)$ $\text{says}_{v(y)}$.

We now define a function (the Tarski jump) $T : S^\omega \to S$ as follows. Fix some arbitrary valuation, v.

$T : S^\omega \to S$

Ts is the function that maps the Gödel number of ϕ to the proposition $\{w \mid w, s, v \Vdash \phi\}$.

We are now in a position to define a set of sequences of semantic value functions:

- $C_0 = S^\omega$
- $C_{n+1} := \{t \in S^\omega \mid \forall n \in \omega, \exists s \in C_n$ such that $t(n) = Ts\}$

Theorem 5.1. *Let $n \geq 0$, and let ϕ be a theorem of T_n. Then $w, s, v \Vdash \phi$ for every $s \in C_n$.*

5.2 Moderate disquotation

Here we give a formalised proof of the argument mentioned in section 3. We introduce a saying connecticate $Say(u, P)$ which takes both a term and a sentence to produce a sentence, and a binary relation $Utt(u, S)$ which tells us when an utterance u is an utterance of a sentence S. To reduce the complexity of the following argument, however, we shall simply assume that quantification involving the variable u is restricted to utterances of the sentence L. Thus $\exists u\phi$ is just short for $\exists x(Utt(x, L) \wedge \phi)$ and $\forall u\phi$ is short for $\forall x(Utt(x, L) \rightarrow \phi)$. (These shorthands should be born in mind when it is noted that something follows from something else in quantificational logic.) Let Q be short for the conclusion $\neg\exists u(\exists P(Say(u, P) \wedge P) \wedge \forall P(Say(u, P) \rightarrow P))$. The instance of moderate disquotation we are interested in states that if an utterance of 'no utterance of L says something true and nothing false' says anything at all it says that no utterance of L says something true and nothing false. Since L just is the sentence 'no utterance of L says something true and nothing false' we can formalise this, employing the preceding shorthands, as follows:

$\forall u(\exists PSay(u, P) \rightarrow Say(u, Q))$

The argument then proceeds as follows

1. $\forall u((\exists P(Say(u, P) \wedge P) \wedge \forall P(Say(u, P) \rightarrow P)) \rightarrow \exists PSay(u, P))$. (Theorem of the logic of propositional quantification inside scope of $\forall u$.)

2. $\forall u(\exists P Say(u, P) \rightarrow Say(u, Q))$. (Premise).

3. $\forall u((\exists P(Say(u, P) \wedge P) \wedge \forall P(Say(u, P) \rightarrow P)) \rightarrow Say(u, Q))$. (From 1 and 2 by transitivity of conditional.)

4. $\forall u(\forall P(Say(u, P) \rightarrow P) \rightarrow (Say(u, Q) \rightarrow Q))$ (universal instantation inside scope of $\forall u$.)

5. $\forall u((\exists P(Say(u, P) \wedge P) \wedge \forall P(Say(u, P) \rightarrow P)) \rightarrow Q)$. (From 3 and 4 by propositional logic inside scope of $\forall u$.)

6. $(\exists u \exists P(Say(u, P) \wedge P) \wedge \forall P(Say(u, P) \rightarrow P)) \rightarrow Q$. (From 5 by quantificational logic.)

7. $\neg Q \rightarrow Q$. (From 6 by definition of Q.)

8. Q. (7 by propositional logic.)

References

[1] M. & T. Williamson Andjelkovic. Truth, falsity, and borderline cases. *Philosophical Topics*, 28:211–43., 2000.

[2] Andrew Bacon, John Hawthorne, and Gabriel Uzquiano. Higher-order free logic and the prior-kaplan paradox. *Canadian Journal of Philosophy*, 46(4-5):493–541, 2016.

[3] T. Burge. Semantical paradox. *The Journal of Philosophy*, 76(4):169–198, 1979.

[4] Herman Cappelen and Ernest Lepore. *Insensitive semantics: A defense of semantic minimalism and speech act pluralism*. John Wiley & Sons, 2008.

[5] Donald Davidson. Radical interpretation. *Dialectica*, 27(3-4):313–328, 1973.

[6] Cian Dorr. To be f is to be g. *Philosophical Perspectives*, forthcoming.

[7] Cian Dorr and John Hawthorne. Semantic plasticity and speech reports. *Philosophical Review*, 123:281–338, 2014.

[8] H. Field. *Saving truth from paradox*. Oxford University Press, USA, 2008.

[9] Hartry H Field. Mental representation. *Erkenntnis*, 13(1):9–61, 1978.

[10] H. Gaifman. Pointers to truth. *The Journal of Philosophy*, 89(5):223–261, 1992.

[11] Michael Glanzberg. A contextual-hierarchical approach to truth and the liar paradox. *Journal of Philosophical Logic*, 33(1):27–88, 2004.

[12] Laurence Goldstein. A consistent way with paradox. *Philosophical Studies*, 144(3):377–389, 2009.

[13] H Paul Grice. Utterer's meaning and intention. *The philosophical review*, pages 147–177, 1969.

[14] J. Hawthorne. Epistemicism and semantic plasticity. *Metaphysical essays*, 1(9):185–211, 2006.

[15] Allen Hazen. Contra buridanum. *Canadian journal of philosophy*, 17(4):875–880, 1987.

[16] DK Lewis. *Convention: a philosophical study*. Harvard University Press Cambridge, 1969.

[17] C. Parsons. The liar paradox. *Journal of Philosophical Logic*, 3(4):381–412, 1974.

[18] A. Prior. On a family of paradoxes. *Notre Dame Journal of Formal Logic*, 2(1):16–32, 1961.

[19] A.N. Prior. *Objects of thought*. Oxford, 1971.

[20] K. Simmons. *Universality and the liar: An essay on truth and the diagonal argument*. Cambridge Univ Pr, 1993.

[21] Scott Soames. *Beyond rigidity: The unfinished semantic agenda of naming and necessity*. Oxford University Press, 2002.

[22] Alan Weir. Token relativism and the liar. *Analysis*, 60(266):156–170, 2000.

[23] CH Whiteley. Let epimenides lie! *Analysis*, 19(1):23–24, 1958.

[24] Bruno Whittle. Self-referential propositions. *Synthese*, pages 1–15, forthcoming.

[25] T. Williamson. Indefinite extensibility. *Grazer Philosophische Studien*, 55:1–24, 1998.

[26] T. Williamson. *Knowledge and its Limits*. Oxford University PRess, 2000.

[27] Timothy Williamson. Everything. *Philosophical Perspectives*, 17(1):415–465, 2003.

[28] Timothy Williamson. *Modal Logic as Metaphysics*. Oxford University Press, 2013.

[29] Timothy Williamson. Semantic paradoxes and abductive methodology. In Brad Armour-Garb, editor, *The Relevance of the Liar*. Oxford University Press, Oxford, Forthcoming.

[30] Elia Zardini. Truth and what is said. *Philosophical Perspectives*, 22(1):545–574, 2008.