Does 'True's Expressive Role Preclude Deflationary Davidsonian Semantics?

Steven Gross
Johns Hopkins University

Introduction

Can one combine Davidsonian semantics with a deflationary conception of truth? A common worry has been that Davidsonian semantics apparently aims to explain meaning, or meanings, in terms of truth, while deflationism holds that the role of truth-talk is wholly expressive, never explanatory. But Michael Williams (1999) argues that this appearance is deceptive: Davidsonian semantics, *contra* Davidson himself, does *not* require truth-talk to play an explanatory role. If this is right, an incompatibilist must produce alternative grounds for doubt. Accordingly, Claire Horisk (2007) would have us shift our attention to 'true's *expressive* role. She argues that, in any event, *it* disqualifies deflationary accounts—at least extant varieties—from combination with Davidsonian semantics: the deflationary accounts' characterizations of the mechanisms by which they can fulfill their expressive function are incompatible with features Davidsonian semantics requires of them. She argues, in particular, that this is so for Quine's disquotationalism, Horwich's minimalism, and Brandom's prosententialism.[1]

In what follows, I argue that Horisk fails to establish her claim in all three cases. This involves clarifying Quine's understanding of a purely referential occurrence; explaining how Davidsonians can avail themselves of a syntactic treatment of lexical ambiguity; and correcting a common misreading of Brandom (answering along the way an objection of Künne (2003) as well). Horisk's piece is of interest in part for the spotlight it shines on these various matters. It is also of interest for a further reason. Although its specific arguments do not succeed, Horisk's paper invites us to consider a more general question concerning what it would be to combine Davidsonian semantics with this or that deflationary conception truth. I conclude with some discussion of that.

Background

I will assume familiarity with both Davidsonian semantics and deflationary approaches to truth. But a brief reminder of the basic ideas will be useful.

Davidson aims to illuminate the concept, or phenomenon, of meaning by asking (i) what knowledge would suffice to put one in a position to understand the speech of another, and (ii) what evidence sufficiently distant from the concepts to be illuminated could in principle ground such knowledge. He answers: knowledge of an appropriate truth-theory for the speaker's language grounded in what sentences the speaker holds true, or prefers true, in what circumstances. A truth-theory allows the derivation, from finite axioms, of theorems of the form 'S is true in L iff p' (*modulo*

---

[1] Bar-On et al. (2000) raise the standard worry. Their fn. 3 supplies references to others. Kölbel (2001) also replies that truth plays no explanatory role in Davidsonian semantics. For their various versions of deflationism, see, e.g., Quine (1970), Horwich (1998a), and Brandom (1994). Davidson (1984) contains his seminal papers on semantics. Davidson's animadversions on deflationism can be found, e.g., in his (1997, 2000, 2005).

context-sensitivity—see below) for all sentences of the target language L, where 'p' is replaced by a sentence that can be said to interpret the target sentence structurally described by 'S.' A Davidsonian "radical interpreter" confirms such a theory in application to some speaker if, while thus interpreting the speaker's sentences, she can also attribute to the speaker attitudes that, given what sentences the speaker holds or prefers true in what circumstances, plausibly optimize her rationality and possession of true beliefs.[2]

There are various versions of deflationism and various attempts to characterize what they have in common (e.g., Stoljar and Damnjanovic, 2010; Armour-Garb & Beall, 2005; Bar-On and Simmons, 2006; Burgess and Burgess, 2011). Williams' characterization, which Horisk deploys as well, will serve our purpose. On this view, deflationists maintain that the function of truth-talk is wholly expressive, never explanatory. Its expressive role is to enable us to endorse or reject indirectly what we cannot endorse or reject directly—for example, because we don't know what was said ("What Joan said, whatever it was, is true—I trust her"), or because the relevant claims are too numerous ("Every claim of the form 'P or not P' is true"). But this exhausts its function. In particular, truth is not a "substantive" property in the sense of one that plays any explanatory role (though of course 'true' may appear in explanations in its expressive capacity). Versions of deflationism—Williams mentions Quine's disquotationalism, Horwich's minimalism, and Brandom's prosententialism—then differ in their accounts of the mechanisms by which truth-talk fulfills this function. We turn to these differences below.

Davidson himself, for various reasons, rejected deflationism (1997, 2000, 2005). While he shared deflationism's hostility to traditional attempts to define 'true' or analyze truth in putatively more basic terms, he concluded only that 'true' is indefinable and truth unanalyzable, not that it is insubstantial in a way that precludes its having an explanatory role. The burden of Williams' paper is to argue that, in any event, nothing in Davidsonian semantics requires truth to play an explanatory role. In a truth-theory, the role of truth-talk is not to explain meaning, or meanings, but to specify truth-conditions for sentences of a language antecedently unknown to a radical interpreter (and thus not available to her for direct endorsement or rejection). Its role in truth-theories is thus expressive. Truth-talk is likewise expressive as used in characterizing the goals of the radical interpreter (thus, as it occurs in Davidson's broader theory of meaning, as opposed to in just the truth-theories that the broader theory of meaning deploys). For example, to say that a radical interpreter should *ceteris paribus* optimize possession of true beliefs is to say *inter alia* that *ceteris paribus* she shouldn't ascribe the belief that snow is white if snow is not white, the belief that there's a bear next to the tree if there isn't, that 2+2=5 if it doesn't, etc.[3]

---

[2] While my discussion, following Williams and Horisk, focuses on Davidsonians who accept both of these answers, I remark below that my main points apply just as much with regard to neo-Davidsonians who reject Davidson's restriction of the semantically relevant evidence to what's available to a radical interpreter.

[3] Williams does not discuss the role of truth-talk in the statement of the radical interpreter's *evidence*—specifically, the ascriptions of holding and preferring *true*. But

Horisk does not challenge Williams on this score. Rather, she argues that 'true's *expressive* role in Davidsonian semantics itself precludes deflationism, or at least extant versions thereof. More specifically, she argues that there are two features required for 'true' to fulfill its expressive role in a Davidsonian truth-theory. First, 'S is true' on the left-hand-side of a T-theorem must refer to a sentence in L.[4] After all, the point of a truth-theory within Davidsonian semantics, at least in part, is to associate truth-conditions with sentences of the target language. How could that be accomplished if sentences of the target language were not picked out by the T-theorems intended to specify this association? Second, the target sentences must be picked out or described by their physical and syntactic properties, not their semantic properties. For truth-theories are supposed to associate truth-conditions with antecedently *uninterpreted* sentences, ones a radical interpreter does not, or cannot take herself to, already understand. Having underscored these features, Horisk argues that utilizing Quine's disquotationalism would violate the first requirement, while utilizing Horwich's minimalism or Brandom's prosententialism would run afoul of the second. The three versions of deflationism Williams mentions thus cannot be combined with Davidsonian semantics—nor, she claims, can any other extant version.

I now argue that in each case Horisk in fact fails to establish her conclusion.

Quine's Disquotationalism

There is some dispute concerning just what Quine's disquotationalist view is and indeed whether he counts as a deflationist (cf. fn. 5 below). But it suffices to again work with Williams' and Horisk's conception. According to Williams, Quine's view has it that

> the meaning of 'true' is fixed by certain logical equivalences. Thus:
>
> > 'All that glisters is not gold' is true if and only if all that glisters is not gold; 'France is hexagonal' is true if and only if France is hexagonal ... and so on.

---

here we see 'true' being used to characterize *indirectly* the endorsements of the subject. The ascriptions are vindicated by the radical interpreter's success, partially on their basis, in ascribing a truth-theory in accord with the canons governing radical interpretation. And the truth-theory supplies the "empirical information" (see below) required to extend our truth-talk to this foreign (or not presumed antecedently understood) tongue.

[4] It can sound odd to require that a *sentence* refer. Frege did indeed thus assimilate sentences to names (albeit with sentences referring to truth-values, not to bits of language *a la* Horisk). But this is not a widely embraced aspect of his views. (Dummett (1973, p. 184) famously labels it a "gratuitous blunder.") It might be thought that Horisk must mean that a *constituent* of the sentence must refer to a sentence of L. But her discussion of Quine (see below) suggests that she does not want this to suffice, assuming a quote-name like 'Snow is white' is a constituent in the relevant sense of ''Snow is white' is true.' Perhaps the feature is best read as requiring that a constituent that refers to a sentence in L occur purely referentially, in the Quinean sense discussed below.

> Appending 'is true' to a quoted sentence is just like canceling the quotation marks ("disquotation"). For this approach, the core use of 'true' involves cases where a sentence is referred to by its quotation-mark name. But given additional empirical information—for example, that a certain sentence expresses Fermat's Last Theorem—we can easily ascribe truth to sentences picked out in other ways. (1999, p. 546)[5]

Horisk's objection to adjoining Quine's disquotationalism to Davidsonian semantics turns on Quine's elaboration of 'true's disquotational effect. Quotes *typically* surround sentential contexts that are not "purely referential." A sentential context is purely referential if and only if one can substitute *salva veritate* expressions flanking a true identity statement. So, 'Cicero' occurs in a purely referential sentential context (so, occurs purely referentially) in

> Cicero was a Roman senator

because one can substitute *salva veritate* such co-referential expressions as 'Tully'—co-referential, because

> Cicero = Tully.

But 'Cicero' does not occur purely referentially in

> 'Cicero' begins with the letter 'c.'

In such cases, the truth of the sentence turns on how things are with the referent of the quote-name—i.e., the name inside the quotes—not on how things are with the referent of the name inside the quotes. But, although quotes *typically* surround sentential contexts that are not purely referential, predication of 'is true' produces exceptions. In

> 'Cicero was a Roman senator' is true,

'Cicero' *does* occur purely referentially, and the truth of the sentence *does* depend on how things are, or were, with its referent. Indeed,

> 'Cicero was a Roman senator' is true iff Cicero was a Roman senator.

Thus 'is true' has the effect of "cancellation of linguistic reference." (Quine, 1970, p. 12) That is, it has the effect of canceling reference *to* terms, rendering relevant instead—so far as the sentence's truth is concerned—the reference *of* those terms.

Horisk then objects as follows. On Quine's view, the truth predicate in such sentences cancels linguistic reference, rendering the context purely referential. So, names

[5] Though we may allow Williams some license, Quine officially might jib at talk of fixing meaning.

inside quotes to which 'is true' is appended function to refer to their usual referents. Moreover, it follows from this cancellation, that the quote-name does not refer to a sentence—nor does anything else (what else would do the trick?). But then, if we conjoin Quine's disquotationalism with Davidsonian semantics, we must maintain that there is *no* reference to target sentences in T-theorems—which violates the first feature required for 'true' to fulfill its expressive function in Davidsonian semantics.

But this argument is not sound. *Contra* Horisk, it does not follow from 'true's rendering 'Cicero's context purely referential that ''Cicero was a Roman senator'' is not functioning to refer to 'Cicero was a Roman senator.' For it's consistent with a context *within* a quote-name's being purely referential that the context *of* the quote-name itself *also* be purely referential. Indicated graphically, the struck-through context and the underlined context

'~~Cicero was a Roman senator~~' is true

which differ only in whether they include the quotation marks, might *both* be purely referential. To see this, we need only apply Quine's test for pure referentiality to the context in which the quote-name itself occurs. Can one substitute *salva veritate* co-referential terms for ''Cicero was a Roman senator'' in

'Cicero was a Roman senator' is true

and thus in

'Cicero was a Roman senator' is true iff Cicero was a Roman senator?

Yes, one can. Suppose, for example, we arbitrarily name the sentence 'Bob,' so that

Bob = 'Cicero was a Roman senator.'

Substituting 'Bob' in the sentences above to yield

Bob is true

and

Bob is true iff Cicero was a Roman senator

does not alter truth-value. So, the quote-name occurs purely referentially in such sentences. It thus refers to the quoted name. At least, Horisk must agree that it does on Quine's view, since this is the test she deploys in her objection.

What of Quine's remark that it is precisely the role of 'true' to cancel linguistic reference? How can we reconcile this with the result of applying Quine's own test—viz., that there is after all reference to a linguistic item? There are two main options. First, we

might grant that Quine commits himself to a contradiction.[6] Of course, it's not advisable for a Davidsonian to conjoin an inconsistent view to her own. But equally obviously she can just disown this remark of Quine's, taking only what she needs. At least she can so long as treating Quine as a deflationist, or utilizing Quine's disquotationalism for deflationist purposes, does not require embracing the further remark. And it does not: what is crucial for Williams is just that instances of the equivalence schema "fix the meaning" of 'true.'

The second option is more charitable to Quine. We can take his talk of canceling linguistic reference as simply expressing his acceptance of the equivalence schema. 'True' doesn't cancel linguistic reference in the sense of rendering the quote-name's context non-referential; rather, it cancels linguistic reference only in that it has the effect, when appended to a sentence's quote-name, of yielding a sentence equivalent to the sentence quoted. There's then no contradiction. We need only take care not to read Quine's talk of cancellation too flat-footedly (a warning suggested by the passage quoted in fn. 4 above).

It should be noted that, in fact, typical T-theorems of a Davidsonian truth-theory for a natural language will not be instances of a disquotationalist equivalence schema. They thus will not be instances of what the disquotationalist considers, in Williams' words, the "core use" of 'true.' (1999, p. 546) First, natural language sentences are typically context-sensitive—tense alone yields this. Straightforward disquotation does not guarantee an equivalent sentence in such cases. To ensure equivalence, one must add some device that ensures that the quoted sentence is understood as it is in fact used out of quotes on that occasion.

Second, T-theorems stated in a language other than the object language will not be disquotational. Indeed, the position of a radical interpreter is stipulated not to include knowledge of the target language. So, even if the target language appears *phonologically*

---

[6] Indeed, it seems he not only *commits* himself to contradictory statements (if we take the "cancellation" remark flat-footedly), he *makes* them—at least if to mention a sentence is to refer to it. Quine writes that:

> … the truth predicate serves, as it were, to point through the sentence to the reality; it serves as a reminder that though sentences are mentioned, reality is still the whole point. (1970, p. 11)

—*contra* Horisk's remark that:

> Generally [on Quine's treatment], a sentence of the form '*s* is true' does not mention the sentence *s*. Furthermore, Quine's treatment gives no reason to suppose that a sentence of the form '*s* is true' refers to *s*; it does not refer to *s* by mentioning it, and there is no other obvious mechanism of reference to *s*. (2007, p. 542)

Perhaps, though, one might distinguish here "genuine" and merely "formal" mentioning *a la* Price (2004, p. 290, fn. 5).

or *orthographically* identical to the radical interpreter's, she cannot presume it to be *semantically* identical: this can only be concluded on the basis of radical interpretation. Thus, even T-theorems of the form "'S' is true-in-L iff S' would be grounded, not in competence with quotation marks, 'true', 'iff', and 'S', but in the radical interpreter's complex holistic empirical justification.

Put otherwise, Quinean disquotation, unlike Davidsonian semantics, concerns context-insensitive sentences we already understand.[7] This means that a Davidsonian deflationist who would exploit Quinean disquotationalism must have something to say about the extension of truth-talk to such cases.[8] But it's not clear what in-principle obstacle this poses. In the case of foreign sentences, we can note that arguably such an extension is precisely what the radical interpreter's deployment of 'true' in ascribed truth-theories provides: the truth-theory supplies the "additional empirical information [that enables us to] ascribe truth" to sentences not (presumed) antecedently understood (adapting the quote from Williams 1999, p. 546, cited above—cf. Williams 2007).

There is also a third reason T-theorems do not involve disquotation: their left-hand-sides do not contain quote-names of sentences, but rather structural descriptions. This figures in our response to Horisk's remarks on Horwich.

Horwich's Minimalism

Paul Horwich argues, *contra* disquotationalists, that 'true' applies fundamentally to propositions. In particular, the meaning of 'true' is determined by our disposition to accept sentences of the form: the proposition that p is true iff p. One can also apply the truth predicate to utterances or sentence-types; but, to do so, one must "individuate [… them] semantically as well as physically." (Horwich 1998a, p. 133)[9] This is because expression-types that are of the same "syntactic form (i.e., physical character)" (Horwich 1998a, p. 100) can be associated with different meanings, as with 'bank' (one of Horwich's examples).

Horisk, on this basis, rules minimalist deflationism unsuitable for Davidsonian semantics. Treating ambiguity this way—*individuating* terms semantically—runs afoul of the second feature Davidsonian semantics requires of 'true's expressive role—viz., that

---

[7] Davidson (1994, pp. 83-4) emphasizes the requirement of understanding in the course of arguing that Quine is not in fact a deflationist. For a discussion of context-sensitivity and disquotationalist deflationism, see Heck (2004).

[8] Indeed, this is the case for *any* deflationist who would exploit Quinean disquotationalism—as, for example, Field (1994, 2001), a main proponent, is well aware. (Note that a *general* objection maintaining that *any* attempt at such extension must fail would target, not in particular the conjoining of disquotationalist deflationism to Davidsonian semantics, but simply disquotationalist deflationism itself.)

[9] Horwich (1998b, p. 104, fn. 1) remarks that "[a]rguably the truth predicate is ambiguous: standing most often for a property of propositions, but sometimes for a corresponding property of utterances," and Horisk follows the orthographic convention Horwich introduces to distinguish the two cases. I will dispense with this. But note that deploying such a convention, as she herself does, parallels what my reply to Horisk below suggests Davidsonians may do to accommodate lexical ambiguity more generally.

target sentences be *picked out or described* by their physical and syntactic properties, not their semantic properties.

Now, ambiguity is a phenomenon Davidsonian semantics must accommodate anyway: if it cannot, it has a serious problem independently of any attempt to combine it with deflationism. On the other hand, if it *can* be accommodated, then Horisk's objection only has bite if the accommodation cannot be combined with Horwich's minimalism. Note that this accommodation needn't preserve Horwich's *own* treatment of ambiguity. A Davidsonian who endorses Horwich's minimalism needn't follow Horwich on every point. Indeed, she *cannot*, since Horwich develops his overall position in part by explicitly rejecting Davidsonian semantics.[10] I will reply to Horisk at first without reference to Horwich's own view of ambiguity. But then I'll suggest that minimal tinkering can perhaps bring his remarks in line with Horisk's requirements.

A common way to accommodate ambiguity in Davidsonian semantics is to insist that T-theories apply to target-language expressions individuated at a grain and in a way appropriate to semantics. For example, the orthographic representation 'Flying planes can be dangerous' is structurally ambiguous. A plausible T-theory will utilize, not this orthographic representation, but rather two representations, distinguished by different syntactic features. Similarly for lexical ambiguity: a plausible T-theory will utilize, not the orthographic representation 'John is at the bank,' but two representations, distinguished syntactically by the appearance of distinct lexical items—if you will, 'bank_1' and 'bank_2,' to which semantic values are assigned by distinct axioms. (Cf., e.g., Gillon 1990 and Larson and Segal 1995.) This requires a sufficiently fine conception of syntax. Though there is a coarser conception according to which 'Dogs run' and 'Cats walk' are syntactically identical, the finer-grained conception relevant to constructing plausible T-theories has them differ in virtue of containing different lexical items: the rest of their trees are the same, but the labels of their terminal nodes differ.

It is consistent with this syntactic treatment of ambiguity, including *lexical* ambiguity, to allow that there is a sense in which semantic considerations play a role in the *justification of the ascription* of a T-theory. A radical interpreter presumably incorporates ambiguity into an ascribed T-theory in this way because doing so appropriately optimizes true beliefs and rational behavior. Competitor T-theories that assign 'bank' a single semantic value via a single axiom would yield less charitable interpretations. In this sense, we might say that radical interpreters individuate expressions in part based on the semantic consequences of doing so. But this of course involves no *antecedent* knowledge of the target expressions' semantic properties. Moreover, we are talking here of an *epistemic* sense of 'individuation.' It concerns what *grounds* the radical interpreter's individuation. *Metaphysical* individuation, on the other hand, concerns what makes it the case that things are distinct or what distinguishes them—not the *basis* for our distinguishing them.

It might be objected that this distinction between the epistemic and the metaphysical collapses in the context of Davidson's broader theory of meaning, since, on his view, the semantic facts (up to indeterminacy) are what they are *in virtue* of a radical

[10] See, e.g., Horwich (1998a, pp. 132-3; 1998b, pp. 71-4). Williams (2007) discusses his disagreements with Horwich.

interpreter's being able to so interpret the speaker—similarly, let us suppose, for these syntactic facts concerning lexical individuation. But though Davidson does *identify* the semantic facts with what radical interpretation can yield, the objection, it my opinion, overstates Davidson's "interpretivism" in suggesting that what *makes it the case* that the semantic facts are as they are is what radical intepretation yields. It is consistent with the tight connection Davidson draws to allow, for example, that non-semantically characterized (actual and possible) *patterns of use* determine the semantic facts—or even to reject the cogency of this 'what makes it the case' question.[11] In any event, even if words *are* typed as they are *in virtue of* the fact that a radical interpreter would so type them, it simply doesn't follow that the expressions mentioned on a T-theorem's left-hand-side are there "picked out or described" semantically. That would be so if the left-hand-sides' structural descriptions incorporated such descriptions as, say, 'the word that is satisfied by financial institutions.' But they don't.

To be sure, T-theories that treat ambiguous words as distinct lexical items *ipso facto* treat them as so individuated as to have distinct meanings. That is indeed the whole point; it cannot be an objection that they do so. Again, it doesn't follow that the lexical items are not individuated syntactically. The expression types '$bank_1$' and '$bank_2$' are syntactically distinct, however tokens may be phonologically, orthographically, or otherwise realized (even if some tokens of different types are not *physically* distinct); and they can be "picked out" by this syntactic difference. It is true that '$bank_1$' and '$bank_2$' *arbitrarily* designate the types that they do: reversing the labels yields a notational variant. But, again, it does not follow, at least not obviously, that one can only pick out or describe the types semantically, even if the arbitrary convention is itself *established* by reference to semantics. Of course questions remain concerning what determines (metaphysically and epistemically) what type a particular token is a token of, but that's another matter.

Horisk does not take up and so of course does not offer any objection to this common way of handling ambiguity. Nor is it obvious that there is anything awry with a Davidsonian *minimalist* deflationist availing herself of it. Indeed, it would seem that Horwich could avail himself of something like it as well. On Horwich's non-Davidsonian conception of meaning, meaning is determined by a use-regularity, what best explains the overall use of the expression-type. For example, his defense of minimalism rests in large part on his claim that a disposition to accept instances of the propositional equivalence schema best explains the overall use of 'true.' As for ambiguity, he accommodates it by noting that in such cases the simplest explanation of the overall use of a *phonological* type involves positing *sub*types with distinct regularities of use. (1998b, p. 81) Positing such subtypes is like positing syntactically distinct lexical items, in that in both cases one

[11] Patterns of use could determine the semantics facts without being identical to them and without the converse entailment: the semantic facts could supervene on use facts without being reducible to them. Horwich (1998b, e.g., p. 5), *contra* Davidson, maintains that semantic facts *can* be reduced to use facts. (He puts the point in terms of properties. My talk of facts is intended as a place-holder neutral among competing metaphysical positions. For what it's worth, Davidson himself, though he does not admit facts into his ontology, likewise indulges in talk of semantic facts.)

is individuating types at a finer grain, so at a further level of abstraction. Thus we may read Horwich's treatment as not really in conflict with Horisk's required feature after all. For Horwich too can say that he individuates lexical items *syntactically*, not semantically—at least under the appropriate construal of syntax. Horwich himself *identifies* syntactic form with physical character (as quoted in this section's first paragraph and by Horisk as well). But there is a conception of syntax available to him that is not so restricted. Indeed, the phonological types to which he adverts are already an abstraction from physical characteristics. Horisk should not have any objection to this: her second feature adverts to "physical and syntactic properties" (2007, pp. 536, 540, 543-4, 547-8, 550, and 556) in a way that strongly suggests that she does not consider the latter a subset of the former. Moreover, there is even a way to preserve Horwich's talk of individuation in terms of *semantic* properties. It might suffice for Horwich's purposes, when he says that sentence-types must be "individuated semantically as well as physically" (1998a, p. 133), that they be *epistemically* individuated semantically and/or so individuated as to yield distinct types when in fact they have distinct meanings.

Brandom's Prosententialism

Brandom differs from both Quine and Horwich in maintaining that, despite appearances to the contrary, 'true' is not even a predicate. On his view, it is a *prosentence-forming operator*. A prosentence, on Horisk's explanation, is one:

> that inherits its content from an anaphoric antecedent, just as many pronouns inherit their content from anaphoric antecedents. To illustrate, suppose Ciara says 'The grass is long', Katrina says 'It needs to be mown', and Connor says 'That is true'. Katrina's pronoun 'it' inherits its content from its anaphoric antecedent, the noun phrase 'the grass'. Similarly, Connor's prosentence 'That is true' inherits its content from its anaphoric antecedent, the sentence 'It needs to be mown'. (2007, pp. 544-5)

Horisk argues that the second feature required for 'true' to play its required expressive role in Davidsonian semantics—again, that target sentences be picked out or described by their physical and syntactic properties, not their semantic properties—precludes adoption of Brandom's prosententialism. For sentences picked out in the required way have their truth-conditions contingently, whereas T-theorems construed in Brandomian terms would be necessary. This modal mismatch arises because, in an instance of an equivalence schema (say, ''Snow is white' is true if and only if snow is white'), the anaphoric antecedent of the clause containing the truth-term is of the same type as the token on the right-hand-side. Since the prosentence *inherits its content* from its antecedent, it is *identical* in content to it. But then it is identical in content to the token on the other side of the biconditional. And a biconditional flanked by synonymous clauses is necessary.[12]

[12] Horisk addresses at some length, drawing upon Rumfitt (2001), replies that would deny that T-theorems are contingent. The basic idea is that, even if languages are individuated semantically, some T-theorems will be contingent—viz., those for strictly *apriori*

A first reply to Horisk redeploys observations made above. In the case of T-theorems, it is not generally true that the token on a T-theorem's right-hand-side is of the same type as the structurally described target sentence: first, because of the deviations required to accommodate context-sensitivity; and, second, because the meta-language may differ from the object-language. Suppose, however, that the target language is in fact the radical interpreter's (though she cannot know this *qua* radical interpreter) and that the target sentence in question is context-*in*sensitive. Even if Horisk's worry does not arise *in general* for Davidsonian T-theorems, it arguably does in *these* cases; and that suffices to cause trouble.[13]

The more fundamental reply to Horisk is that her characterization of Brandom's view is based on a common misreading, unfortunately abetted by some of Brandom's own locutions but clearly not what he intends.[14] On Brandom's prosententialism, the content of a prosentence is not *exhausted* by that of its antecedent. So, although it *inherits content* from its antecedent, it's not the case that *all* of the prosentence's content is thus inherited, as the phrase 'inherits its content'—not only used by Horisk, but unfortunately found in Brandom as well—would suggest.[15] If Brandom's view *were* as Horisk

---

statements. Strictly *apriori* statements are those such that anyone who understands them are, on that account, in a position know them—i.e., those whose *truth* is entailed by an interpretive semantic theory. While the T-theorems' for sentences expressing such statements will have necessary left-hand-sides (since they will be entailed in all worlds by the theory), some such statements (and thus the T-theorems' right-hand-sides) will be contingent, rendering the T-theorem as a whole contingent.

I will not discuss this issue here, but rather will show that, even if we grant the premise that T-theorems are contingent, we can reply to Horisk's claim that the expressive role of 'true' precludes combining Davidsonian semantics and Brandomian prosententialism.

[13] In Davidson's view, what language someone speaks is indeterminate. (Cf., e.g., Davidson, 1979, pp. 239-40; 1994, p. 82.) So, talk here of supposing the radical interpreter and the target subject speak the *same* language would have to mean something like: they use phonologically very similar expressions that *can* be interpreted more-or-less homophonically, context-sensitivity aside. I am trying, however, to keep indeterminacy in abeyance. In any event, the supposition is to Horisk's advantage. –That said, even if the languages *do* differ, the first reply at best undermines Horisk's *argument* that Brandom's view commits him to the necessity of T-theorems (at least in the case of context-insensitive sentences); it does not show that the view is *not* so committed. The next reply does.

[14] The misreading is found as well in, for example, Armour-Garb and Beall (2005) and Burgess and Burgess (2011).

[15] The phrase 'inherits its content' and close cousins appear five times in *Making It Explicit* (1994, pp. 304, 305, 327, 493, and 568—thank you Google Books). Most, perhaps all, *can* be read as *not* entailing identity of content. In some cases, however, it arguably requires a stretch. For instance, Brandom (1994, p. 305) writes of "treating the whole expression 'that is true' as a prosentence anaphorically referring to a sentence

characterizes it, it would be open to objection *independently* of any attempt to combine it with Davidsonian semantics. For suppose John utters 'It's raining' and Mary utters 'What that tall guy said is true.' It would seem Brandom would have to maintain that the content of Mary's utterance is simply that it's raining. And this is objectionable because *prime facie*—and on Brandom's own views of content and inferential status (see below)—it would seem that the utterance's content contains more, involving being a tall guy and saying. (Cf. Burgess and Burgess, 2011, p. 41.) But Brandom is not guilty of the charge. The objection is based on a misreading.[16]

Though Brandom does use the phrase 'inherit its content' and says, for example, that ''Snow is white' is true' has the "same semantic content" as its antecedent, he immediately after the last remark adds that:

> [t]he prosentence differs from its antecedent in explicitly acknowledging its dependence upon an antecedent—as 'She stopped' differs from 'Mary stopped' when the pronoun has some token of the type 'Mary' as its antecedent. Otherwise the lazy uses are purely redundant. (1994, p. 302)[17]

Their content thus differs precisely in that the *contents* of such sentences as ''Snow is white' is true,' unlike that of 'Snow is white,' in part concern linguistic tokens. And it's not just that Brandom *says* this: there is nothing in his view that precludes him from doing so.

Moreover, Brandom is well aware of the differing modal profiles of such sentences. This is clear from remarks he makes in developing his parallel deflationary account of reference.

> This account of Tarskian contexts in which 'refers' and its cognates appear [according to which base clauses are so paraphrased that 'refers' occurs only within an anaphorically indirect description] respects the different modal status of
>
> The term 'Leibniz' denotes Leibniz,
>
> which is only contingently true, and

---

tokening from which it inherits its content." Here, one would have to read 'it' as referring to the expression 'that is true' and 'its' as referring to the antecedent.

[16] There *are* maneuvers available to a view that *does* maintain identity of content. One might distinguish the presuppositional and assertoric content of an utterance, or maintain that an utterance of the sentence would effect several distinct assertions. But Brandom needn't go any such route, even if his other commitments leave room for them (something that is not altogether clear).

[17] In an attached footnote, he even takes the last claim back, foreshadowing aspects of lazy pronouns' inferential role that become important for him later. (Incidentally, Brandom (p.c.) confirms that his 'Snow is white is true' on p. 302—p. 328 too—is just a typo, *pace* Künne 2003, p. 84.)

> Leibniz is Leibniz,
>
> which is necessarily true. For the possibility that the first claim is not true can be understood in terms of its paraphrase as the existence of a possible world *w* such that
>
> The one referred to as 'Leibniz' in *w* is not Leibniz,
>
> That is, is not the one we refer to in our own world as 'Leibniz'. The explicit relativization of the indirect description to a possible world simply specifies which world its antecedent tokens are found in. The candidate antecedents of
>
> the one referred to as 'Leibniz' in *w*
>
> are tokenings of the type 'Leibniz' that are uttered in *w*. The anaphoric approach accordingly has room for what has been thought of as the contingency of word-world semantic relations, although it is not based on such relations. (1994, p. 318)

Not only do these remarks take up these modal differences, they attempt to illuminate them in light of the anaphoric deflationary account. Returning to Brandom's prosentential theory of 'true,' the analogous point is that the contingency of

> 'Snow is white' is true iff snow is white,

in contrast to the necessity of

> Snow is white iff snow is white,

reflects the fact that, in different possible worlds, the prosentence would anaphorically depend on different antecedent tokenings.

If Brandom does not identify the content of a prosentence with that of its antecedent, then Horisk's objection to combining prosententialism with Davidson semantics vanishes. There is a worry that persists, however. I have suggested that the prosentence contains "more" content than its antecedent. This of course leaves in place that it contains *at least* the antecedent's content, which can seem objectionable enough. Thus Künne (2003, p. 86) invites us to consider someone who has heard of Goldbach's Conjecture but knows nothing that would distinguish it from various other theorems about prime numbers. Suppose this person reads a headline announcing its proof. How, Künne asks, could her claim that Goldbach's Conjecture is true have the same content as (or, we can add, have even in common the content of) the claim that every even number greater than 2 is the sum of two primes? This, to be sure, is an objection to Brandom's prosententialism itself, not specifically to combining it with Davidsonian semantics; nor

need a reply to *Horisk* rebut it. But taking it up will clarify Brandom's position and our reply to Horisk.

The crucial point here is that one must understand Brandom's talk of content in terms of his larger enterprise. For him, it is a shorthand for more official talk of inferential relations of incompatibility, commitment preservation, and entitlement preservation. 'What the newspaper says is true' differs in content from 'Every even number greater than 2 is the sum of two primes' because asserting the former, but not the latter, commits one to the existence of a newspaper: it thus contains "more" content. But someone who asserts the former *does* commit herself, given the chain of antecedents, to the latter—and entitles others to it as well—even if, in her ignorance of prime number theory, she doesn't realize this. Her assertion inherits the inferential relations of its antecedent; it's in this sense that it inherits its content. Brandom (2010, p. 337) says he can talk of inferential commitments "without committing myself to identifying and individuating *concepts*, or *conceptual contents* (though sometimes, when it seems harmless, I do speak with the vulgar and use such expressions)." It seems that this is a case in which it has not proven harmless.[18]

It might be suggested that, while this clarifies Brandom's content-talk, the question remains whether anything in *Davidson's* views—whatever might be the case with Brandom—might require him, were he to adopt prosententialism, to identify the contents of prosentences and their antecedents. So put, this is not yet a developed objection, but we can point out in any event that at least the two most obvious places one might look for support in fact don't offer any. First, some ascribe to Davidson the view that the right-hand-side of a T-theorem *gives the meaning* of the sentence targeted on the left-hand-side (e.g., Lepore and Ludwig 2005). Might this force a prosententialist Davidsonian to endorse the synonymy of prosentences and their antecedents? Actually, this ascription is disputed (Ebbs 2012). But, even if accepted, it raises no problem here. One needs to consider the T-theorems for sentences *containing* 'true,' which, on a prosententialist approach, would presumably assign the sentence a prosententialist structure. Such a T-theorem would perforce be in some manner contextualized. For example, it might take the form: for utterances u of sentence-type S, for all contextual parameters c_1, …, c_n, if $F(c_1, \ldots, c_n)$, then u is true iff $G(c_1, \ldots, c_n)$. Among the contextual

---

[18] Künne (2003, p. 86) raises a second objection. Brandom (2002, p. 106) says that, to understand a sentence containing 'true,' "one must process the noun phrase to determine what sentence tokening (or class of such tokenings) it picks out as anaphoric antecedent(s)." Künne objects that this standard needn't be met, rhetorically asking whether one can assert with understanding

> Hegel's most notorious remark about truth is true, but I have no idea what that remark is

only by falsifying its second conjunct. But the reply is that knowing-what comes in degrees (perhaps 'know what' is context-sensitive). The speaker knows what the antecedent is, to a degree sufficient for Brandom's criterion, in knowing that it is Hegel's most notorious remark about truth.

parameters would be those that determine which sentence is the prosentence's antecedent. Such a T-theorem would not render a prosentence-token and its antecedent synonymous. (Cf. Gross 2005a on context-sensitivity and meaning-theoretic extensions of Davidsonian truth-theories.) Second, there is Davidson's (1968) deployment of a same-saying relation in dealing with indirect speech. Might he be committed to maintaining that prosentences and their antecedents same-say one another? But, again, this raises no problem. Same-saying, for Davidson, amounts to being relevantly similar, where relevance no doubt varies with context. (Cf. Davidson 1999.) Davidson can allow the relevant similarity of prosentences and their antecedents—for instance, in what other attitudes the speaker who uses them might be, on that account, reasonably taken to have—without running afoul of Horisk's requirements.

Conclusion

Horisk's challenge fruitfully forces us to ask whether the *details* of developed deflationary accounts mesh with the *details* of Davidsonian semantic theorizing. I have argued that, so far as her specific objections go, Davidsonian deflationists remain unscathed. But that's not to say there are no more devils in the details. I have certainly not argued that deflationism and Davidsonian semantics *are* compatible. But, by removing further alleged obstacles, I have given further reason to think they may be. I have also not argued that deflationist Davidsonian semantics is an attractive position. There are of course standard worries for deflationism and for Davidsonian semantics taken individually.[19] Combining them does not remove the worries. I have argued only that combining them has not been shown to *add* to the worries. Those who find each individually attractive—and who think the standard worries can be answered—have reason to avail themselves of the combination.

I would like to conclude by deflecting a question our discussion might suggest, one that can seem more general and prior to the specific points we have taken up—viz., just what *is* it to combine this or that deflationist view with Davidson's? In various places, we have noted that of course it cannot mean conjoining *everything* Quine, or Horwich, or Brandom says with what Davidson says. To add to previous instances: Davidson, unlike Horwich, eschews propositions as explanatorily otiose; and Brandom, unlike Davidson, rejects compositionality. Indeed, might one, on Horisk's behalf, press this as an objection: until we have more clarity on what counts as such a combination, perhaps we are not in a position to assess—or to endorse—one?

But 'what counts as such a combination?' is simply not the important question. The important question is whether one can develop a deflationism that meshes with Davidsonian semantics. Such a development could take this or that element from whatever extant view it wished, so long as the result works and receives support. Perhaps there's an argument, for example, that a radical interpreter would assign prosentential logical forms to subjects' target-language uses of truth-talk. If so, it just wouldn't matter whether the position in other respects was Brandomian. We began, after all, by taking this attitude towards Davidson's own views, since he rejects deflationism.

---

[19] In Gross (2001, 2005b, and forthcoming), I discuss in particular worries for truth-conditional semantics arising from pervasive context-sensitivity.

Indeed, there are other aspects of Davidson's views that a *neo*-Davidsonian deflationist may well wish to drop. For example, cognitivist neo-Davidsonians (e.g., Larson and Segal 1995), who ascribe cognition, or embodiment, of a Davidson-style truth-theory as part of the larger psychologized project of empirical linguistics, reject Davidson's restriction of the semantically relevant evidence to what's available to a radical interpreter. (Cf. Gross 2012.) Construing the truth-talk in their ascriptions in a deflationary way would allow such semanticists to deploy a truth-conditional framework without representational commitments. As far as I can see, Horisk's worries do not arise for such neo-Davidsonians either.[20]

Bibliography

Armour-Garb, B., and J. Beall (eds.). 2005. *Deflationary Truth* (Open Court).

Bar-On, D., C. Horisk, and W. Lycan. 2000. "Deflationism, Meaning and Truth-Conditions," *Philosophical Studies* **101**, pp. 1-28.

Bar-On, D., and K. Simmons 2006. "Deflationism," in *The Oxford Handbook of the Philosophy of Language*, E. Lepore and B. Smith (eds.) (Oxford).

Brandom, R. 1994. *Making It Explicit* (Harvard).

Brandom, R. 2002. "Expressive vs. Explanatory Deflationism about Truth," in *What Is Truth?*, R. Schantz (ed.) (de Gruyter), pp. 103-19.

Brandom, R. 2010. "Reply to Jerry Fodor and Ernie Lepore's 'Brandom Beleaguered,'" in *Reading Brandom*, B. Weiss and J. Wanderer (eds.) (Routledge), pp. 332-7.

Burgess, A., and J. Burgess. 2011. *Truth* (Princeton).

Davidson, D. 1968. "On Saying That," *Synthese* **19**, pp. 130-46. Reprinted in Davidson (1984).

Davidson, D. 1979. "The Inscrutability of Reference," *The Southwestern Journal of Philosophy* **10**, pp. 7-19. Cited as reprinted in Davidson (1984).

Davidson, D. 1984. *Inquiries into Truth and Interpretation* (Oxford).

Davidson, D. 1994. "What is Quine's View of Truth?" *Inquiry* **37**, pp. 437-40. Cited as reprinted in his *Truth, Language, and History* (Oxford, 2005).

Davidson, D. 1997. "The Folly of Trying to Define Truth," *Journal of Philosophy* **94**, pp. 263-78.

Davidson, D. 1999. "Reply to Reinaldo Elugardo," in *Donald Davidson: Truth, Meaning and Knowledge*, U. Zeglen (ed.) (Routledge), pp. 119-20.

Davidson, D. 2000. "Truth Rehabilitated," in *Rorty and his Critics*, R. Brandom (ed.) (Blackwell), pp. 65-74.

Davidson, D. 2005. *Truth and Predication* (Harvard).

Dummett, M. 1973. *Frege: Philosophy of Language* (Harvard).

Ebbs, G. 2012. "Davidson's Explication of Meaning," in *Donald Davidson on Truth,*

---

[20] For helpful conversation or correspondence, I thank Claire Horisk, Alexis Burgess, Michael Williams, Robert Brandom, Michael Glanzberg, and Gary Ebbs. Thanks also to this volume's anonymous referees and to attendees of my lab meeting, of a seminar Michael Williams and I co-taught, and of the 2012 meeting of the Society for Exact Philosophy.

Meaning, and the Mental, G. Preyer (ed.) (Oxford), pp. 76-104.
Field, H. 1994. "Deflationist Views of Meaning and Content," *Mind* **103**, pp. 249-85.
Field, H. 2001. "Postscript" [to Field 1994] in his *Truth and the Absence of Fact* (Oxford).
Gillon, B. 1990. "Truth Theoretical Semantics and Ambiguity," *Analysis* **50**, pp.178-82.
Gross, S. 2001. *Essays on Linguistic Context-Sensitivity and its Philosophical Significance* (Routledge).
Gross, S. 2005a. "The Biconditional Doctrine: *Contra* Kölbel on a 'Dogma' of Davidsonian Semantics," *Erkenntnis* **62**, pp. 189-210.
Gross, S. 2005b. "Context-Sensitive Truth-Theoretic Accounts of Semantic Competence," *Mind & Language* **20**, pp. 68-102.
Gross, S. 2012. "Davidson, First Person Authority, and the Evidence for Semantics," in *Donald Davidson on Truth, Meaning, and the Mental*, G. Preyer (ed.) (Oxford), pp. 228-48.
Gross, S. forthcoming. "(Descriptive) Externalism in Semantics," *The Routledge Handbook of Semantics*, N. Riemer (ed.) (Routldge).
Heck, R. 2004. "Truth and Disquotation," *Synthese* **142**, pp. 317-52.
Horisk, C. 2007. "The Expressive Role of Truth in Semantics," *The Philosophical Quarterly* **57**, pp. 535-57.
Horwich, P. 1998a. *Truth*, 2nd ed. (Oxford).
Horwich, P. 1998b. *Meaning* (Oxford).
Kölbel, M. 2001. "Two Dogmas of Davidsonian Semantics," *Journal of Philosophy* **98**, pp. 613-35.
Künne, W. 2003. *Conceptions of Truth* (Oxford).
Larson, R., and G. Segal 1995. *Knowledge of Meaning* (MIT).
Lepore, E., and K. Ludwig 2005. *Donald Davidson: Meaning, Truth, Language, and Reality* (Oxford).
Price, H. 2004. "Naturalism without Representation," in *Naturalism in Question*, M. De Caro and D. Macarthur (eds.) (Harvard), pp. 71-88 and 290-1.
Quine, W. 1953. "Notes on the Theory of Reference," in his *From a Logical Point of View*, cited according to the 2nd ed. (1980), pp. 130-8.
Quine, W. 1970. *Philosophy of Logic* (Prentice-Hall).
Rumfitt, I. 2001. "Semantic Theory and Necessary Truth," *Synthese* **126**, pp. 283-324.
Stoljar, D., and N. Damnjanovic. 2010. "The Deflationary Theory of Truth," *The Stanford Encyclopedia of Philosophy* (Winter 2010 Edition), E. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2010/entries/truth-deflationary/>.
Williams, M. 1999. "Meaning and Deflationary Truth," *Journal of Philosophy* **96**, pp. 545-64.
Williams, M. 2007. "Meaning, Truth and Normativity," in *Truth and Speech Acts*, D. Greimann and G. Siegwart (eds.) (Routledge), pp. 377-95.