

Journal for the History of Analytical Philosophy

Volume 6, Number 9

Editor in Chief

Marcus Rossberg, University of Connecticut

Editorial Board

Annalisa Coliva, UC Irvine

Henry Jackman, York University

Kevin C. Klement, University of Massachusetts

Consuelo Preti, The College of New Jersey

Anthony Skelton, Western University

Mark Textor, King's College London

Audrey Yap, University of Victoria

Richard Zach, University of Calgary

Editor for Special Issues

Sandra Lapointe, McMaster University

Review Editors

Sean Morris, Metropolitan State University of Denver

Sanford Shieh, Wesleyan University

Design and Layout

Daniel Harris, Hunter College

Kevin C. Klement, University of Massachusetts

jhaponline.org

Truth, Meaning, and Interpretation: A Reconsideration of Davidson's Program

Arpy Khatchirian

On a common reading of Davidson, the motivation for his proposal that a meaning theory is to take the form of a truth theory is at least partly guided by concern with the ends and means of interpretation. At the same time, the consensus seems to be that this proposal faces a particularly stubborn justificatory burden. The aim of this paper is twofold: first, to suggest that there is a promising route to discharging this burden, albeit one that is visible only once we shift our attention away from the so-called 'problem of interpretation'; second, to make the case that, contrary to initial appearances, the line of justification offered here gives us a plausible interpretation of Davidson's own goals.

© 2018 Arpy Khatchirian

Truth, Meaning, and Interpretation: A Reconsideration of Davidson's Program

Arpy Khatchirian

1. Introduction

We linguistically competent beings have both the capacity to make meaningful utterances and the capacity to understand the meaningful utterances of others. This paper concerns Davidson's famous proposal that a meaning theory for a language should center on the construction of a compositional truth theory, and the way in which distinguishing between these two capacities can help shed new light on this proposal.

It is widely assumed that a meaning theory for a language is a theory of understanding for that language, despite considerable disagreement about what agreeing to this slogan commits one to. My interest here is in the following question: should a theory of meaning for a language be primarily taken to concern a speaker's ability to speak the language or her ability to understand other speakers' utterances in the language? But need we choose? A tempting, and, I believe, generally accepted thought, is that 'knowledge of' a language L simply involves the capacity to understand each of the potential infinity of sentences in L, and it is this understanding (together with many other abilities that are not in the province of a theory of understanding for L) on which speakers of L draw both in making meaningful utterances in L and in understanding the meaningful utterances of other speakers of L.

Attractive as it may seem, this picture will not help make sense of Davidson's own conception of the task of a meaning theory for a language, given Davidson's idiolectal approach to language (see 'A Nice Derangement of Epitaphs,' and 'The Social

Aspect of Language' in [Davidson 2005b](#)). The primary notion, for Davidson, is that of the language spoken by a speaker on a particular occasion, and we need not speak the same language as another speaker in order to be able to understand her utterances. Thus, my making meaningful utterances and my understanding the meaningful utterances of others cannot be taken to be two aspects of a capacity we might call 'my competence in a language L', even when other speakers turn out to be speaking the language I am speaking. If this is right, it does make sense to ask which of these two capacities Davidson is primarily concerned to illuminate, in proposing that a meaning theory take the form of a truth theory. And if we, as I do, share Davidson's position on the primacy of the idiolect, it also makes sense for us to ask which of these two capacities we should take to be the proper concern of a meaning theory.

What makes this question particularly difficult to answer is a long-standing lack of agreement about the sense in which a meaning theory for a language L is supposed to count as a theory of understanding for L, be it the understanding of a speaker or of an interpreter that is in question. In particular, how do the claims entailed by a meaning theory, which ascribe certain features to expressions of L, but do not ascribe any capacities to speakers or interpreters of L, bear on the capacities involved in understanding L?¹

Here, then, are the four questions I am interested in, starting with this last one: First, what is the sense in which a meaning

¹To give a sense of the wide range of positions here, I will mention a couple: at one extreme is the idea that a meaning theory for L is a theory of understanding for L only in the sense that explicit knowledge of the theory would suffice for understanding any arbitrary utterance in L. At another extreme is a view that attributes to competent speakers tacit knowledge of the axioms of an acceptable meaning theory, akin to Chomskyan tacit knowledge of a grammar (see [Larson and Segal 1995](#)). For an illuminating assessment of the different ways of reading the requirement that a meaning theory be a theory of understanding, see [Smith \(1992\)](#) (though as will become clear, I disagree with some aspects of Smith's reading of Davidson there).

theory for L is a theory of understanding for L? Second, where L is the language spoken by a given speaker on a particular occasion, is a meaning theory for L a theory of the speaker's understanding of her own language, or is it a theory of an interpreter's ability to understand the speaker's utterances? Third, can a truth theory for L be said to be a meaning theory for L, and if not, what role should a truth theory be taken to play in a meaning theory? And since my interest here is in making sense of Davidson's own motivation for a truth-conditional approach to meaning theories, and I see this as bound up with his conception of a meaning theory as a theory of understanding, my fourth question is what we should take Davidson's answers to the first three questions to be.

Here are the answers I want to motivate in this paper. First, and with nuances to be introduced in due course, a meaning theory for L is a theory of understanding for L at least in the sense that its main explanandum is a competent user's knowledge of certain semantic facts about sentences of L, where this knowledge is taken to be constitutive of understanding L, and where the meaning theory aims to explain this knowledge as having a compositional basis. My assumption here is that understanding L involves explicit—or, available to consciousness—propositional knowledge of certain semantic facts about sentences of L. This leaves it open whether, and the sense in which, a speaker should be attributed knowledge of the axioms of an acceptable meaning theory for her language.²

²Thus, my claim should be distinguished from the kind of 'cognitivism' defended by, among others, [Larson and Segal \(1995\)](#), [Evans \(1981\)](#), and [Davies \(1987\)](#), who all attribute to a speaker tacit knowledge of the axioms of an acceptable meaning theory and undertake an account of tacit knowledge on which this attribution is explanatorily significant. My claim concerns, not the axioms of a meaning theory, but some of its theorems at the level of sentences, and it attributes to a speaker explicit propositional knowledge of these theorems. I should also mention [Higginbotham \(1989\)](#), who takes (as I do) a speaker's idiolect to be the primary object of study, and who thinks of a truth theory for a speaker's language as at least part of an account of the

Second, where L is the language spoken by a given speaker on a particular occasion, a meaning theory for L is, in the first instance, a theory of the speaker's understanding of her own idiolect, rather than of an interpreter's ability to understand the speaker's utterances. Third, a suitable truth theory for L should itself be taken to be a meaning theory for L, rather than a component of a meaning theory, or something to which a meaning theory makes reference. Putting these three claims together: in proposing that a meaning theory is to take the form of a truth theory, we are proposing to explain a speaker's knowledge of her own language as partly consisting in knowledge of the truth conditions of her sentences. We are not proposing to explain an interpreter's ability to understand the speaker's utterances as partly consisting in his knowledge of the truth conditions of the speaker's sentences (note that this does not involve denying that an interpreter who understands the speaker's utterances has knowledge of their truth conditions).

Finally, and despite evidence that might appear to suggest otherwise, I believe that we have good reasons to think that Davidson would ultimately agree with these answers. My reading provides a plausible motivation for Davidson's proposal to use truth theories as meaning theories, where other interpretations have failed to do so, and it makes sense of otherwise puzzling features of his approach.

speaker's knowledge of her language. Despite these parallels, there are striking differences between our approaches. For the sake of brevity, I will only mention one fundamental difference between our conceptions of the nature and role of knowledge of truth conditions: for Higginbotham, a speaker's knowledge of such facts as that 'Snow is white' is true in her language if and only if snow is white does not amount to much—it is knowledge that the speaker can have even if she lacks knowledge of what 'snow' and 'white' mean. [Higginbotham \(1992\)](#) seems to replace the idiolect with the communal language as object of study, and he goes on to explain a speaker's knowledge of the language as consisting in her knowledge, not of truth conditions, but rather, of what she and others are expected to know about truth and reference simply by virtue of being competent users of the language.

I want to motivate the joint plausibility of these four claims by arguing for the demise of a very different and widely shared conception of the task of a meaning theory for a language L, and of a corresponding reading of Davidson. On this reading, the main task Davidson assigns a meaning theory for L is simply that of specifying the content of knowledge that would suffice to interpret any arbitrary utterance in L. I will argue that this reading fails to provide any plausible justification for the proposal to use truth theories as meaning theories. Thus, I agree with [Soames \(2008\)](#) that, on the prevalent interpretation of Davidson's goals, his project is hopeless. But I will argue that the moral should be, not that Davidson's project is indeed hopeless, but that we have misunderstood, and that Davidson himself has at times misconstrued, the motivation for a truth-conditional approach to meaning.

Here is the structure of the rest of the paper. In Section 2, I argue that, in asking 'What knowledge would suffice for interpretation?', we have been asking the wrong question. In particular, I focus on Ernest Lepore and Kirk Ludwig's proposed improvement on Davidson's account: their proposal to think of a meaning theory for a language L as an explicit statement of what we could know, *about* a compositional truth theory for L, that would put us in a position to interpret any utterance in L. I argue that despite its ingenuity, this proposal ultimately fails to sufficiently motivate a truth-conditional approach. In Section 3, I sketch a different conception of the goal of a meaning theory, and of the way in which a truth theory is supposed to serve as one, and I explain why it is immune to the difficulties plaguing the approach examined in Section 2. My proposal is that using a truth theory as a meaning theory involves explaining a speaker's knowledge of her language in terms of her knowledge of the truth conditions of her sentences. It does not involve giving an account of what knowledge could put us in a position to interpret utterances in the language. Here are some virtues of my proposal, that the justificatory attempts examined in Section 2 do not pos-

sess: (a) it gives substance to the conception of a meaning theory as a theory of understanding, (b) it has an attractive simplicity, and (c) it helps explain, rather than relying on, the oft-invoked distinction between what a truth theory says and what it shows. Of course, these virtues would have been earned only if we can motivate the assumption that a speaker's understanding of her language involves her knowledge of the truth conditions of her sentences. But as we will see, Lepore himself, among others, has done much to reveal the crucial role of knowledge of truth conditions in communication. Finally, I argue that the roots of my proposal can be found in Davidson's own writings, and I explain away apparent evidence to the contrary.

2. We Have Been Asking the Wrong Question

2.1. The problem of interpretation

Davidson typically presents his proposal concerning the form to be taken by a meaning theory as an answer to a certain problem about interpretation. 'Radical Interpretation,' most notably, begins with the following queries:

Kurt utters the words 'Es regnet' and under the right conditions, we know that he has said that it is raining. Having identified his utterance as intentional and linguistic, we are able to go on to interpret his words: we can say what his words, on that occasion, meant. What could we know that would enable us to do this? ([Davidson 1984](#), 125)

Davidson then goes on to generalize these questions, asking, 'What knowledge would serve for interpretation?' ([1984](#), 126). The answer he goes on to give is that knowledge of an appropriately constrained Tarski-style truth theory would do the job. This, I think, is a perfectly plausible answer to the question Davidson asks, if only because there are many ways in which a theory could play a role in interpretation. However, as we will

see, neither the answer, nor the question, helps shed light on the motivation for using truth theories as meaning theories.

First, let me briefly introduce some terminology and further clarify the problem Davidson sets out to solve in such papers as 'Truth and Meaning' and 'Radical Interpretation' (Davidson 1984). A Tarski-style truth theory for a language L is a theory modeled after Tarski's proposed definitions of truth for particular languages (see Tarski 1944, 1956). For our purposes, we can think of such a definition as a system of axioms that finitely and recursively generates, for each sentence of L, a theorem of the form:

(T) ⟨s is true in L if and only if p⟩

where what is to be substituted for 's' is a quotation-name of this sentence in the metalanguage (i.e., the language of the theory), what replaces 'p' is a sentence of the metalanguage, and 'if and only if' is the material biconditional. Following common usage, I will call these sentences 'T-sentences', or 'T-theorems'.

Here are two fairly uncontroversial assumptions I will be making. First, I take the claim that a speaker has, on any given occasion, linguistic abilities spanning infinitely many sentences to be built into Davidson's way of spelling out the 'problem' of interpretation, rather than invoked to explain the possibility of interpretation.³ Second, in asking what knowledge would enable us to interpret any utterance in a given language, Davidson is assuming that there is a core component of the significance of an utterance that derives from a compositional account of the language as a whole (Davidson 1984, 53). What he is asking is, 'What could we know that would put us in a position to determine the literal meaning of each utterance in the language spoken by a speaker on a particular occasion?', for some suitable notion of literal meaning.⁴ While he often uses the term 'inter-

pretation' in the broader sense of interpreting a speaker (which involves attributing content to his attitudes), I will here be mostly be concerned with his notion of 'interpreting an utterance' in the sense of understanding, or determining, its literal meaning.⁵

What, then, is Davidson's answer to the problem of interpretation? Since it involves an appeal to truth theories, it is natural to suppose that this answer is, simply, that knowledge of an acceptable truth theory for a speaker's language would suffice for grasping the literal meaning of the speaker's utterances. Indeed, this is precisely what Davidson seems to be proposing, not just in his early writings, but later on:

[A truth theory] gives the substance of what a knowledgeable interpreter knows which enables him to grasp the meaning of the speaker's utterances. (Davidson 2005a, 52)

A theory of truth for a speaker is a theory of meaning in this sense, that explicit knowledge of the theory would suffice for understanding the utterances of that speaker. (Davidson 2005a, 53)

Of course, it is knowledge of an *acceptable* truth theory that is in question here. What is the relevant notion of acceptability? Acceptability for Tarski was a matter of satisfying his famous Convention T. To satisfy Convention T, a recursive definition of 'is true-in-L' for language L is to entail, for each sentence of L, a T-theorem, where 's' is replaced by a structural description of the sentence, and 'p' is replaced by the sentence itself (if the metalanguage contains L), or by an appropriate translation of this sentence in the metalanguage.

Tarski's goal was to ensure that for each language L, his predicate 'is-true-in-L' picks out all and only the true sentences of L. Davidson, by contrast, wants to put recursive characterizations of truth to use as meaning theories. But this does not mean

of a sentence as uttered on a particular occasion need not correspond to any conventionally determined meaning.

⁵Of course, this is not meant to deny that interpreting an utterance, in the sense I am interested in here, usually involves interpreting a speaker.

³Davidson (1999) directly supports this reading.

⁴On Davidson's idiolectal conception of a language, the 'literal meaning'

that Convention T is irrelevant to Davidson's own goals, even if Davidson wants to spell out the empirical constraints on an acceptable truth theory without relying, as Tarski does, on an *unexplained* notion of translation (see [Davidson 1984](#), 172–73). So when Davidson claims that explicit knowledge of an acceptable truth theory would suffice for understanding a speaker's utterances, we can take him to be claiming that explicit knowledge of a *translational* truth theory would suffice for understanding (where a truth theory is translational if it satisfies Convention T).

But as Davidson himself acknowledges in some of his more guarded moments, knowledge of what is stated by a translational truth theory for L is simply not sufficient for grasping the literal meaning of any arbitrary utterance in L. One reason for this is that we can know what is stated by a translational truth theory without knowing that our knowledge is stated by such a theory.⁶ Unsurprisingly, then, Davidson ends up accepting Foster's (1976) appraisal of the situation: though knowledge of the facts entailed by a translational truth theory for L would not suffice for interpreting utterances in L, knowledge of these facts, together with knowledge that these facts are entailed by such a theory, would suffice for interpreting any arbitrary utterance in L.

2.2. Instrumentalism

If this is Davidson's answer to the question 'What knowledge would suffice for interpretation?', we can plausibly agree with it (though see [Soames 2008](#) for doubts about this). What is not clear, however, is why a truth theory, or the notion of truth itself, is really needed here. Indeed, a number of philosophers have argued that the work done by a truth theory in enabling us to interpret a speaker's utterances can be carried out whether or

⁶As Davidson ends up conceding to John Foster, even when an interpreter 'has a theory that satisfies Convention T, nothing in the theory itself tells him this' ([Davidson 1984](#), 173).

not we take its predicate 'is true' to express any particular notion of truth. For instance, [Williams \(1999\)](#) argues that using a truth theory as a meaning theory for L does not involve explaining meaning facts as consisting in facts about truth conditions. As Williams sees it, a meaning theory is just a 'recursive device for specifying the meaning of every sentence of a given language' ([Williams 1999](#), 553). But why should such a 'device' involve the construction of a truth theory? This question is made all the more urgent by the observation that for any language like ours, a recursive truth theory is simply not an easy thing to construct! Williams answers this question along familiar lines:

In specifying the meaning of sentences of another speaker's language, we associate his sentences with sentences of our own. In doing so, we make use of the truth predicate, which is what lends color to the idea that Davidson explains meaning in terms of truth-conditions. But the use of 'true' in a Davidsonian meaning specification for a particular speaker is expressive, not explanatory. He eschews 'means that' [in target theorems of the form $\langle s \text{ means in } L \text{ that } p \rangle$] in favor of the material biconditional . . . To replace 'means that' with a sentential connective, we need a sentence on the left side . . . This is precisely what 'true' allows us to form. ([Williams 1999](#), 557)

On this picture, a truth theory for L serves as a meaning theory for L only insofar as it recursively generates theorems in which sentences of L are paired with metalanguage sentences that can be used to 'interpret them' (in some appropriate sense). And just as the truth theory is a recursive device for pairing sentences of L with metalanguage sentences in use, its truth predicate is a syntactic device that enables us to generate such pairings. But it seems, then, that 'is true' can play its role as such a device regardless of what it is taken to mean, and arguably, regardless of whether it is taken to mean anything at all. If our goal is simply that of pairing sentences of L with interpretations, the only crucial constraint on our interpretation of 'is true' seems to be

that it be construed as of the right syntactic type⁷—substituting it for ‘...’ should turn the following string of symbols into a sentence:

(*) ‘La neige est blanche’ . . . if and only if snow is white.

Is there any reason at all why the device used to effect such pairings of object-language sentences with their interpretations should express a notion—or, any particular notion—of truth?

2.3. Truth theories in radical interpretation

At this point, it may look like we have overlooked one crucial piece of the puzzle, namely, the constraints governing the construction of a truth theory in radical interpretation. It may be the very nature of these constraints that impose further demands on the truth predicate beyond that of recursively generating pairings of object-language sentences with metalanguage sentences in use.

Indeed, Davidson himself cites the empirical character of truth theories in explaining why his approach assigns truth a crucial explanatory role. Consider the following passage:

A theory of truth, viewed as an empirical theory, is tested by its relevant consequences, and these are the T-sentences entailed by the theory. A T-sentence says of a particular speaker that, every time he utters a given sentence, the utterance will be true if and only if certain conditions are satisfied. T-sentences thus have the form and function of natural laws; they are universally quantified biconditionals, and as such are understood to apply counterfactually and to be confirmed by their instances. Thus, a theory of truth is a theory for describing, explaining, understanding, and predicting a basic aspect of verbal behavior. Since the concept of truth is central

⁷Of course, Williams might disagree with this assessment, since he describes ‘is true’ as playing an expressive—albeit not explanatory—role in a Davidsonian semantics, and this would seem to require more than its being of the right syntactic type. But as far as I can tell, Williams does not explain why (or what) more is required of this predicate.

to the theory, we are justified in saying that truth is a crucially important explanatory concept. (Davidson 2005a, 54)

This passage clearly does not sit well with Williams’s depiction of ‘true’ as simply a device for pairing sentences with interpretations. As Davidson sees it, it is our grasp of truth, a concept applicable to our own sentences as well as the sentences of others, that we bring to the task of constructing and testing a truth theory for the language of another speaker.

But is this right? Is the way a truth theory is ‘tested’ in radical interpretation a matter of directly confirming the truth of its T-sentences (construed as generalizations about the conditions under which utterances of particular sentences are, or would be, true)? More generally, do the constraints governing the construction of a truth theory in radical interpretation depend, for their intelligibility, on any particular way of understanding its truth predicate?

There are three constraints we need to make sense of. The first constraint is that of identifying logical form in the speaker’s sentences. This, for Davidson, will involve imposing our logic onto the language we are trying to interpret, where this is assumed to be first order logic. The second constraint is that we maximize truth in the beliefs attributed to the speaker—or, maximize agreement between the speaker and ourselves. And finally, a constraint the importance of which Davidson clearly emphasizes in his later writings, is that we take the speaker’s utterances, in certain basic cases, to be about and true of the very features of the environment that cause them.

Williams argues that none of these constraints makes any explanatory demands on the notion of truth. If this is right, it would make it sufficiently plausible that the methodology of radical interpretation is not tied to any particular way of understanding the truth predicate employed by a Davidsonian truth theory. Williams’s discussion of the first two constraints seems to me to be largely on point, so I will be pretty brief here. While

I also think that he draws the right conclusion about the third constraint, this will require a bit more discussion to convince.

If Davidson is right, the need to read our logic into the speaker's language depends on assuming overall logical consistency on her part. But does it draw on any particular way of understanding truth? As Williams explains, privileging first-order logic 'does not . . . require that we build our logical preferences into our concept of truth' (Williams 1999, 560). I would add that reading first-order logic into the speaker's language does not depend on any particular way of interpreting the predicate 'is true' employed in our truth theory for her language. What forces first-order logic into the language are the formal constraints imposed by the construction of a Tarskian truth theory, rather than any particular way of understanding its truth locution.

How about the need to maximize truth in the speaker's beliefs? Williams argues that despite appearances, the operative principle here is really that of agreement maximization, and therefore, not one that makes any explanatory demands on truth, since it need not even mention truth. But I think a stronger point can be made here. For whether we spell out the relevant principle in terms of agreement maximization or truth maximization is irrelevant: on either construal, it is easy to see there are no explanatory demands on truth. Spelled out in terms of agreement maximization, the principle involves attributing to the speaker beliefs that agree with ours, i.e., attributing to the speaker the belief that p only if we ourselves believe that p . Spelled out in terms of truth maximization, the principle involves attributing, for the most part, true beliefs to the speaker. But this is a matter of attributing to the speaker the belief that p only if p , so talk of truth is only playing a generalizing role here.⁸

⁸Of course, there are infinitely many sentences or beliefs to account for, and some beliefs count for more than others, which is why Davidson goes on to replace the notion of agreement 'maximization' with that of agreement 'optimization' (I thank an anonymous referee for this observation; see 'Truth and Talk,' Davidson 1984, 169.) This, however, does not affect the main point here.

Let us move on to the third constraint, that of taking a speaker's utterances, in certain basic cases, to be about and true of their mutually salient causes in our shared environment. This principle helps generate hypotheses about the interpretation of so-called 'observation sentences' in the speaker's idiolect—sentences the speaker's holding-true of which (or, assent to which) seems to systematically vary with mutually salient changes in our environment. As Davidson sees it, it is here, in the early stages of theory construction, that truth emerges as a crucial explanatory primitive. Contrasting his distal approach to the interpretation of observation sentences with Quine's proximal approach, Davidson assimilates the difference between the two to 'the opposition between a theory of meaning that makes evidence primary, and a theory of meaning that makes truth primary' ('Meaning, Truth, and Evidence,' Davidson 2005b, 58).

But what exactly is the difference between the proximal approach and the distal approach? Why favor the distal approach over the proximal, and what does this have to do with the significance accorded to truth? Here is how Davidson spells out the difference between the two approaches:

On the proximal theory, . . . [observation] sentences have the same meaning if they have the same stimulus meaning—if the same patterns of stimulation prompt assent and dissent . . . The distal theory, on the other hand, depends primarily on shared causes which are salient for speaker and interpreter, learner and teacher. Meanings are shared when identical events, objects or situations cause or would cause assent and dissent. As a radical interpreter I correlate verbal responses of a speaker with changes in the environment. Inferring a causal relation, I then translate those verbal responses with a sentence of my own that the same changes in the environment cause me to accept or reject. (Davidson 2005b, 54)

Note that the notion of truth does not explicitly figure in this summary of the contrast between the two approaches. The main difference is that what is taken to matter in interpretation is, in the one case, the matching of sentences alike in stimulus meanings

(where stimuli are construed as proximal), and in the other case, the matching of sentences our assent to which is caused by the same mutually salient conditions in the environment.

But as Davidson further explains, the reason why we should favor a distal approach is that tying the meaning of observation sentences to their proximal causes cannot make sense of what it would take for such sentences to be true. To see this, suppose we follow Quine and tie the meaning of an observation sentence to its proximal causes. How should we think of truth for such a sentence? Should we tie truth to proximal causes, or should we tie it to salient features of the environment? Choosing the first option involves taking the truth of a speaker's observation sentence to be a matter of how things are with her proximal stimulations, rather than with the shared external world she appears to be talking about. But this is unacceptable, not just because it deprives the truth of our sentences of any connection to the shared environment we take ourselves to be thinking and talking about, but also, because it severs the connection between interpersonal sameness of meaning and sameness of truth conditions. This is why Quine himself chooses the second option, and ties the truth of observation sentences to how things are with the 'real external objects' (Quine 1981, 181) that impinge on the speaker's nerve endings. But as Davidson observes, this option is equally problematic if meaning remains tied to proximal causes, since it leaves open the possibility of attributing global error to a speaker's basic beliefs about her perceptual environment.

The upshot is that the proximal approach to meaning leaves us with no plausible conception of truth—in Davidson's words, it is an approach on which meaning or translation is 'in danger of losing track of truth' ('Pursuit of the Concept of Truth,' Davidson 2005b, 78). This is what motivates Davidson's distal approach and helps explain why he describes this approach as one that 'makes truth primary.' But on closer look, it is not at all clear that the distal approach's ability to yield a notion of meaning that keeps track of truth itself makes any explanatory

demands on truth. From a perspective that takes translation to be the relevant task, our goal, in adopting a distal approach, is to yield a notion of translation that allows for truth to be preserved by translation. But this is a goal we can meet whether or not we think of truth as intelligible prior to translation, and thus whether or not we attribute to the interpreter grasp of a primitive intersubjective notion of truth. We could adopt a distal approach to the translation of observation sentences in terms of mutually salient causes of our utterances of such sentences. Once translation is under way, we could then secure the sought-for connection between meaning and truth by defining truth for the languages of others in such a way that it is preserved by translation.⁹

The conclusion of this section is that the important advantages of Davidson's distal approach to the interpretation of observation sentences over Quine's proximal approach need not be spelled out as differences in their conceptions of truth, or in any explanatory demands imposed on this notion. If this is right, the methodology of radical interpretation does not help explain Davidson's need for a substantive, explanatory notion of truth.¹⁰

⁹In other words, the connection between translation and truth can be secured whether we invoke an interpersonal notion of truth to constrain translation, or appeal to an independently intelligible notion of translation to constrain our definitions of truth for the languages of others.

¹⁰Ramberg (1989) would strongly disagree with this assessment. He argues that underplaying the differences between radical interpretation and radical translation only serves to obscure crucial advantages of Davidson's truth-centered approach over Quine's. I cannot do justice to Ramberg's intricate argument here, but I do hope that the following brief remarks will give a sense of where I think it falls short. As Ramberg sees it, one crucial advantage of Davidsonian radical interpretation over Quinean translation is that it avoids the threat of reified meanings. Quinean translation leads to such reification insofar as involves assigning stimulus meanings to a speaker's observation sentences, then finding sentences of ours that have the same, or relevantly similar, stimulus meanings. By contrast, all a radical interpreter needs to do is directly specify the mutually salient conditions that cause the speaker to hold a sentence true. On the assumption that these are conditions under which

This brings us back to Williams's conception of the role of the truth predicate in a Davidsonian truth theory as simply that of a syntactic device for pairing object-language sentences with metalanguage interpretations. Williams himself does not hesitate to describe a truth theory *as* a meaning theory. After all, a meaning theory is for him just a recursive device for generating pairings of sentences of a language with metalanguage sentences in use that interpret them. But as we have seen, this is not quite how Davidson sees things in 'Radical Interpretation.' A meaning theory as he conceives of it there is not a mere recursive device, if it is that at all. Rather, it states something we could *know* that would suffice for interpretation. Could we reconcile this condition on a meaning theory with the conception of a truth theory as just a recursive device for generating interpretations?

2.4. Lepore and Ludwig's proposal

This is precisely what Ernest Lepore and Kirk Ludwig seem to be trying to do. They accept Davidson's requirement that a meaning theory for a given language is to spell out the content of knowledge that would suffice for interpreting utterances in it. Only, they argue that Davidson and many of his commentators

the speaker's sentence is in fact true, the interpreter has thereby produced the truth conditions of the sentence. There is no further task of matching a sentence of hers with these conditions. While I am sympathetic to Ramberg's worries here, I think that this does not quite pin down a crucial advantage of interpretation over translation. The main problem here is that it is not clear that the notion of translation really needs to be explained in terms of the matching of stimulus meanings. More generally, it is not clear that making sense of translation has to involve, as Ramberg puts it, 'reification, thinking of meanings as something to be captured by, by given independently, of the sentences we use' (1989, 67). Notice also that once we adopt a distal approach, what Ramberg says about interpretation, we could say about translation: in describing the relevant causes of a speaker's holding-true of an observation sentence, we have thereby produced a sentence of ours that translates the speaker's—there's no further task of finding a sentence of ours that matches those conditions.

err in describing the knowledge in question as knowledge of a truth theory, or even, as including knowledge of a truth theory (i.e., of the facts entailed by a truth theory). Their proposal is to think of a meaning theory for L as an explicit statement of what we could know, *about* a compositional truth theory for L, that would put us in a position to interpret any arbitrary utterance in L.

Here is their suggested outline for, as they call it, 'an explicit compositional meaning theory stated in terms of knowledge of an interpretive truth theory':

- [1] Every instance of the following schema is true:
For all speakers X , times t, s for X at t in L means that p iff it is canonically provable on the basis of the axioms of an interpretive truth theory T for L that for all speakers X , times t, s for X at t is true in L iff p .
- [2] \mathcal{T} is an interpretive truth theory for L whose axioms are . . .
- [3] Axiom . . . of \mathcal{T} means that . . .
Axiom . . . of \mathcal{T} means that . . .
. . .
- [4] A canonical proof in \mathcal{T} is . . . (Lepore and Ludwig 2005, 120–21)

An 'interpretive' truth theory, in Lepore and Ludwig's sense, is a theory whose axioms are interpretive—where, roughly, an axiom is interpretive if it states the semantic contribution of an object-language expression using a metalanguage expression that translates it (see Lepore and Ludwig 2005, 72). The point of the appeal to a canonical proof procedure is to generate T-theorems that only draw on the content of the axioms, thus guaranteeing interpretiveness at the level of T-theorems.

The goal here is to finitely generate, for each sentence of the object-language, a uniquely identifiable T-sentence, from which a corresponding M-sentence can be derived, where an M-sentence is of the form:

(M) ⟨For all speakers X , times t, s (for X at t in L) means that p)
(or, simplifying, ⟨ s in L means that p ⟩)

Assuming we can successfully construct truth theories that have the required properties, we can plausibly agree with Lepore and Ludwig that knowledge of the sort of explicit ‘meaning theory’ they outline here (what I will refer to as an ‘LL-theory’) would indeed suffice for interpretation. But in what way does this vindicate the pursuit of LL-theories, or explain their significance?¹¹

As Lepore and Ludwig insist, what puts us in a position to understand a language L is not knowledge of any facts entailed by a truth theory for L , but rather, the sort of knowledge they describe, knowledge *about* an interpretive truth theory for L . But what this does not sufficiently explain is why an account of what knowledge could put us in a position to interpret utterances in L should center on the construction of a truth theory. In particular, why should it not center on the construction of a compositional translation theory, whose theorems explicitly pair each sentence of L with a suitable translation in our own language?

2.5. Truth theories versus translation theories

As is often stressed, knowing what a sentence means goes beyond knowing that it is equivalent in meaning to some other sentence. In the same vein, knowing what is stated by an acceptable translation theory from a language L onto another language

¹¹A different kind of attempt to exploit Davidsonian truth theories in the context of theories entailing M-sentences can be found in Kölbel (2001). Unlike an LL-theory, a Kölbel-style meaning theory for L itself includes a Davidsonian truth theory—in the sense of entailing everything that a truth theory entails—but in addition also entails an M-theorem for each sentence of L . To achieve this result, Kölbel proposes that we simply add to our logic an inference rule that permits us to derive an M-sentence from each canonically derived T-sentence. I do not think that this proposal is more plausible than Lepore and Ludwig’s, since Kölbel’s inference rule is not valid: it can be used to derive false M-sentences from axioms that are true but not interpretive.

L' (that is, a translation theory pairing each sentence of L with a sentence of L') need not, by itself, put us in a position to understand either language. However, when L' is our own language, knowledge of (what is stated by) an acceptable translation theory from L onto L' , together with knowledge that L' is our own language, would put us in a position to understand any arbitrary utterance in L .

But this means that with respect to their ability to state the content of knowledge that would put us in a position to interpret a language L , an interpretive truth theory for L and a translation theory from L into our own language are on a par. Neither sort of theory states facts knowledge of which would suffice for interpreting utterances in L . But as we have seen, knowledge of certain facts about an interpretive truth theory would suffice for interpreting utterances in L , as would knowledge of a translation theory from L into L' , along with knowledge that L' is our own language. Of course, there is a difference between a truth theory for L and a translation theory from L into L' : as Davidson explains (1984, 129), a translation theory from L into L' does not work for any interpreter. But why require a theory that would work for any interpreter? This Davidson does not explain.¹²

Thus, if nothing more is said about the work done by a truth theory here, we would have to agree with Scott Soames’s skeptical assessment:

The only role played by knowledge of that which is stated by a translational truth theory (with canonical theorems) is that of allowing the agent to identify a unique canonical claim in which S is paired with a certain content, which is recognized . . . to be the content expressed by a translation of S , and hence by S itself. Neither the truth of this canonical claim, nor the fact that it states the truth-conditions of S , plays any role in deriving the interpretation of S . All it does is supply a translation, which could be supplied

¹²My contention is that an adequate explanation would only go to show that what Davidson is really interested in here is the task of capturing a speaker’s knowledge of her own language.

just as well in other ways—with or without the notion of truth. (Soames 2008, 11)

Lepore and Ludwig (2011, 273) do in fact agree with Soames's claim that a truth theory need not be true, and its predicate 'is true' need not be interpreted as a truth predicate, for the theory to do its work in enabling us to understand sentences of a given language. But despite this, they maintain that an interpretive truth theory does more than indirectly supply translations. What more does it do? More importantly, since a translation theory *directly and explicitly* supplies translations, while a truth theory does so only *implicitly and indirectly*, why bother using truth theories? Here is Lepore and Ludwig's answer:

Proofs of the canonical theorems *exhibit* how parts of sentences, in virtue of their meanings, contribute to fixing the truth-conditions of these sentences, by way of using terms the same in meaning. We see exhibited in the proof the semantic structure of the sentence and how it fixes truth-conditions. This is not what the proof says, but it can be culled from the proof. Someone in possession of such a theory and appropriate knowledge of it is in a position to understand the compositional structure of the language. That is more than being able to pair object language sentences with metalanguage ones that translate them. (Lepore and Ludwig 2011, 273)

So unlike a translation theory, a truth theory for L appears to be playing a dual role here. First, it generates interpretations of utterances of L by indirectly supplying translations of sentences of L onto sentences of the metalanguage. Second, it sheds light on the compositional structure of L by entailing statements of the conditions under which each sentence of L is true on the basis of assumptions governing its structure and the contributions of finitely many parts.

2.6. Response

I have two points to make in response to this. First, if we accept Lepore and Ludwig's proposal, then it does, after all, matter how

we interpret the truth locution that figures in our truth theory. This does not matter as far as the goal of generating interpretations is concerned. But unless we take 'is true' to express a notion of truth, it is not clear how the truth theory is supposed to shed any light on how the truth conditions of whole sentences depend on features of their parts. Moreover, the T-theorems have to express facts, or *true* claims, about the truth conditions of sentences of the language, in order for their derivability from the axioms to shed light on any of this. It is not clear to what extent Lepore and Ludwig would agree with this. On the one hand, the following passage appears to suggest that they would agree with it:

[A]s Davidson observed, if the theory is true, the predicate has the extension of the truth predicate. And if the predicate is the truth predicate and the axioms meet Convention A and the primitive expressions are not semantically defective, then the theory is true, and we can retrieve more information from the theory than what is expressed in its theorems of form:

s means that *p*. (Lepore and Ludwig 2011, 273)

On the other hand, they seem to be divided on the question whether a truth theory's ability to shed light on the compositional structure of a language really does depend on its yielding true theorems: Lepore seems inclined to agree with this, while Ludwig maintains that a truth theory need not be true in order to play its role in an LL-theory (see Lepore and Ludwig 2005, 137 n).

Second, it is not clear how the two tasks Lepore and Ludwig assign a meaning theory (stating the content of knowledge that would suffice for interpretation, and giving a compositional account of the language) are connected. In the following passage, Davidson seems to be explaining the connection between these two tasks:

[W]e can interpret a particular sentence provided we know a correct theory of truth that deals with the language of the sentence. For

then we know not only the T-sentence for the sentence to be interpreted, but we also ‘know’ the T-sentences for all other sentences; and, of course, all the proofs. Then we would see the place of the sentence in the language as a whole, we would know the role of each significant part of the sentence, and we would know about the logical connections between this sentence and others. (Davidson 1984, 138–39)

Clearly, however, seeing ‘the place of [a] sentence in the language’ in the sense of knowing the facts entailed by a truth theory for the language of this sentence is neither necessary nor sufficient for being in a position to interpret utterances of this sentence.¹³

Of course, Lepore and Ludwig agree with this assessment, and this is part of what motivates their proposal concerning the shape to be taken by a meaning theory. But it is not clear how they propose to construe the connection between the two tasks. They do connect the interpretiveness requirement on a truth theory with its ability to give a compositional account of the language: as they see it, only an interpretive truth theory can be said to give a compositional account of the language, by showing ‘how parts of sentences, in virtue of their meanings, contribute to fixing the truth conditions of these sentences, by way of using terms the same in meaning’ (Lepore and Ludwig 2011, 273). But it is not clear what this comes to. In particular, it is not clear whether a compositional account of a language, as they are thinking of it, really amounts to a compositional account of *meaning* for L, or only to a compositional account of *truth* for L. Does a truth theory satisfying Lepore and Ludwig’s constraints explain (or ‘show,’ or ‘exhibit’) how the meanings of whole sentences depend on features of their parts and their

¹³Not necessary, since I can know (for instance, by being told) that some sentence, as uttered by a given speaker on any occasion, literally means that snow is white, without knowing anything about the rest of the speaker’s language. Not sufficient, as Davidson himself acknowledges, but not just because a truth theory may be true without being translational, but because I may know the facts entailed by a translational truth theory without knowing that my knowledge is entailed by such a theory.

modes of composition, or does it only explain how the truth conditions of whole sentences depend on features of their parts and their modes of composition? If they have the former in mind, it is unclear how this is supposed to work. The interpretiveness of a truth theory’s axioms, along with the restriction to a canonical proof procedure, guarantees that the theory can be used to show what each expression or sentence of the language means. But as Heck (2007) explains, this does not amount to showing how the meanings of sentences *depend on* the meanings of their parts, even if the theory successfully shows how the truth conditions of sentences depend on features of their parts.

In any case, even if the interpretiveness of a truth theory is a requirement on both its ability to give a compositional account of the language and its playing a role in an acceptable LL-theory, this does not, by itself, adequately connect the two tasks. All we are given is a common requirement on fulfilling them; this does not explain what fulfilling the one task has to do with fulfilling the other. Moreover, we cannot even attribute these tasks to one and the same theory: it is the truth theory that is supposed to give us a compositional account of the language, and the LL-theory (of which, it bears emphasizing, the truth theory is not even a component) that is supposed to spell out the content of knowledge that would suffice for interpretation.

3. A Truth Theory as an Account of a Speaker’s Knowledge of Her Own Language

3.1. The proposal

I think we can do better than this. As I already indicated, my starting point is the requirement that a meaning theory for a language should be a theory of understanding for that language. On my reading, this is a substantive constraint on a meaning theory: it involves taking facts about understanding to constrain the shape to be taken by a meaning theory. But how might con-

sideration of what it is to understand a language motivate the proposal that a meaning theory take the form of a truth theory? And how is a truth theory for a language L supposed to help shed light on the capacities involved in understanding L?

If our goal, in constructing a truth theory for L, really does have anything to do with shedding light on understanding, it is a speaker's understanding of her own language, rather than our ability to interpret her utterances, that is the only plausible subject of illumination here. It is only insofar as we are interested in capturing the structure of a speaker's knowledge of her own language that truth theories have any role to play over and beyond that of finitely generating interpretations. If this is right, the main point of a meaning theory for a given speaker's language is not to explain how we do, or how we could, understand a speaker's sentences on the basis of understanding their parts (contra [Lepore and Ludwig 2011](#), 265). It is to explain how the speaker herself understands each of her sentences on the basis of understanding their parts. And the way in which a truth theory does this is by giving a compositional account of what the speaker knows about the conditions under which each sentence of her language is true. Thus, on my reading, using a truth theory as a meaning theory involves explaining a *speaker's* knowledge of the meanings of her sentences in terms of her knowledge of their truth conditions.¹⁴ Here, I mean to include not just the speaker's

¹⁴Contrary to what Lepore and Ludwig argue in [2011](#). Here's their response to the suggestion (in [Soames 2008](#)) that Davidson is trying to explain knowledge of meaning in terms of knowledge of truth conditions:

What would it be to explain "knowledge of meaning"? *Whose* knowledge? What *kind*? It is most natural to interpret the project so described as explaining the knowledge of speakers of a language. But, then, Soames's suggestion would seem to be that Davidson was proposing speakers know their languages by way of knowing a truth theory—how else to interpret "in terms of knowledge of truth conditions," which would seem to be propositional knowledge. Davidson, however, *explicitly* denies this. He is not offering a psychological theory about the mechanism by which speakers understand. He has no commitment about whether "knowledge of mean-

ing" is explicable in terms of propositional knowledge. ([Lepore and Ludwig 2011](#), 269-70)

occurrent knowledge of the truth conditions of those sentences she happens to be using on the relevant occasion, but also, her dispositional knowledge of the conditions under which other sentences of her language would be true if uttered.

One way in which a truth theory for L could help explain the knowledge involved in understanding sentences of L is by entailing statements of the content of this knowledge, on the basis of appropriate axioms governing finitely many semantic primitives and modes of combination. Of course, the connection between the truth theory and the content of explicit knowledge of truth conditions is likely to be more indirect than this, since the truth theory might yield statements of what a speaker of L knows about the truth conditions of her sentences only in conjunction with additional assumptions external to the theory (especially in light of context-sensitivity in the language). However, to simplify exposition, I am here ignoring this complication and taking the requirement to be that the truth theory itself entails statements of what the speaker knows about the conditions under which each of her sentences is true.

This picture avoids many of the difficulties plaguing attempts to explain the point of truth-conditional meaning theories in

ing" is explicable in terms of propositional knowledge. ([Lepore and Ludwig 2011](#), 269-70)

Lepore and Ludwig seem to be assuming that the only sense that can be made of the idea that Davidson is trying to explain knowledge of meaning in terms of knowledge of truth conditions is by saddling him with the following two goals:

1. Explaining a speaker's knowledge of her language as resting on her propositional knowledge of the contents of an acceptable truth theory (in a sense that involves her knowledge of the theory's axioms).
2. Giving an account of the psychological mechanisms by which a speaker understands her language.

It is clear that Davidson rejects both of these characterizations of his goals (for example, see [Davidson 1990](#), 211–12). However, neither of these goals is forced on us by my proposal that in using a truth theory to describe a speaker's competence, we are attributing to the speaker (propositional) knowledge of the T-theorems entailed by this theory, and taking this knowledge to be constitutive of her competence.

terms of the ends and means of interpretation. First, it avoids trivializing the conception of a meaning theory as a theory of understanding. By contrast, it is not clear why a theory knowledge of which could put us in a position to interpret utterances in L should count as a theory of understanding for L. More precisely, it is not clear how, in specifying the content of knowledge that *could* (at least when appropriately supplemented) put us in a position to interpret a speaker, we are shedding light on either our capacity to interpret the speaker, or the speaker's own understanding of her language.

Secondly, the picture I am offering is attractive for its simplicity. Unlike Lepore and Ludwig's account, it does not involve any multiplication of tasks, levels of theory, or levels of explanation within the meaning theory. It also avoids reliance on an unexplained distinction between what a truth theory says and what it shows:¹⁵ as I see it, acceptable T-theorems *state*, rather than merely *show*, what the speaker knows about each of her sentences in understanding them the way she does. In fact, my reading helps explain the special sense in which a T-theory can be used to show what each sentence of a language means: a truth theory can be used to show what each sentence of the speaker's language means insofar as it entails statements of what the speaker knows about each of her sentences (in understanding its parts and their modes of composition). It bears emphasizing that these theorems *state*, rather than merely *show*, what the speaker knows about each of her sentences in knowing its meaning. But what no truth theory itself entails is the claim that its T-theorems have this property. This is the sense in which

¹⁵By contrast, those who try to explain the point of a truth theory in terms of its role in interpretation often rely on a distinction between what a theory *says* and what it *shows*, or between *being* a meaning theory and *serving as* one. For instance, Lepore and Ludwig remark that what a truth theory 'shows it does not say' (2005, 264–65 n). Similarly, John McDowell (1998, 173) has suggested that a Davidsonian truth theory is meant to *serve* as a theory of sense rather than *being* one.

our using one true truth theory rather than another can show what we take sentences of the language to mean, even though neither theory entails statements of what these sentences mean.

But why think that linguistic competence (knowledge of one's language) involves metalinguistic knowledge? And why think that the relevant knowledge is knowledge of truth conditions? Here I will defer to what I take to be some of the most promising attempts to defend these assumptions: see Lepore (1983, 1997), Heck (2006, 2007), and Rumfitt (1995). Each of these authors emphasizes the role that our knowledge of truth conditions plays in explaining both our linguistic acts and our distinctive ways of fulfilling some of the basic purposes of communication (such as the transmission of knowledge). The motivating thought here is that a speaker's capacity to make meaningful utterances is not just a matter of her being disposed to manipulate words and respond to stimuli in such a way as to be interpretable as saying certain things. Rather, it involves the speaker's knowledge of what she is doing with the words she chooses to use.

As Heck puts it, speech is intentional not just under propositional descriptions, but also under verbal descriptions: we correctly take ourselves and others not just to be saying things, but also, to be intentionally using certain words to mean and say these things. The alternative to this is to take ourselves to be 'quasi-telepaths' (to borrow an expression in Heck 2006): as beings who communicate with others, but without any awareness, on our part, of how our using the words we do enables us to communicate. Some may find this picture more attractive, on account of the fact that it imposes fewer demands on linguistic competence. In particular, many of us are reluctant to attribute grasp of metalinguistic concepts (such as those of truth and meaning) to children and ordinary speakers. But it is not clear why we should prefer a less demanding notion of linguistic competence over a more demanding one—provided of course, that there are reasons for any demands we impose. Davidson, for one, clearly did provide such reasons: both thought and language, he argued,

require grasp of the concept of error, or equivalently, grasp of an intersubjective notion of truth (see 'Rational Animals,' 'The Second Person,' and 'The Emergence of Thought,' in [Davidson 2001](#)).

As I see it, the attribution to a speaker of knowledge of the truth conditions of her sentences need only be part of a story about what the speaker knows, or can do, in speaking the way she does: it is the central, compositionally derived core. The rest of the story will involve explaining how the speaker draws on this knowledge in speaking the way she does. Thus, even if we are convinced that linguistic competence crucially involves propositional knowledge of features of our sentences, the plausibility of my further claim that the knowledge in question is knowledge of truth conditions will depend on the extent to which attribution of such knowledge can be shown to do all the explanatory work that attribution of metalinguistic knowledge is meant to do. As I hope is made clear by what I have said so far, the level of explanation I take to be relevant here is that of rational psychology, though this is perfectly compatible with taking the internal articulation of the truth theory to correspond to sub-personal levels of description.

This also helps explain why we should not make much of Davidson's claim that knowledge of truth conditions is *sufficient* for understanding, since, on my reading, we do not start out with a clear enough independent conception of what a speaker's understanding of her language consists in, by reference to which we can judge whether or not knowledge of truth conditions is indeed sufficient for understanding. Rather, the plausibility of the proposal that a speaker's linguistic competence partly consists in her knowledge of truth conditions will have to be judged by reference to the viability and fruitfulness of a conception of understanding as involving knowledge of semantic facts, and the extent to which the attribution to speakers of knowledge of truth conditions can help explain the things they can do with words.

3.2. Textual evidence: speaker and interpreter

Finally, I would like to make the case that my proposal is not only one that Davidson should have endorsed, but also, one that he would have been prepared to accept. Despite Davidson's more explicit preoccupation with the question of what knowledge would suffice for interpretation, there is, I believe, a good case to be made for thinking that it is a speaker's knowledge of his own language that he all along sought to capture by means of a compositional truth theory.

Indeed, once we scratch the surface, even parts of 'Radical Interpretation' help support my interpretation. Consider, again, Davidson's brief discussion in it of the advantages of a truth theory over a translation theory, as far as each theory's ability to specify the content of knowledge that would suffice for interpretation is concerned. Davidson agrees that an interpreter could use a translation theory to interpret another speaker's utterances, but argues that in doing this, the interpreter 'brings to bear two things he knows and that the theory does not state: the fact that the subject language is his own, and his knowledge of how to interpret utterances in his own language' ([Davidson 1984](#), 129–30).

What I am suggesting is that this explanation as to why a truth theory would constitute a better solution to the 'problem of interpretation' than a translation theory, only goes to show that what Davidson is really after here does not primarily concern interpretation. Rather, Davidson's main concern is to give a description of a speaker's linguistic competence—to spell out, in structure-revealing terms, what a competent speaker knows about each of his sentences in understanding it the way he does. Davidson says that a translation theory leaves out 'what we need to know that allows us to interpret our own language' ([Davidson 1984](#), 130). But really, when it comes to our own language (that is, the language we actually now speak), our speaking it, or understanding it, does not involve interpreting our own utter-

ances in it, as Davidson himself acknowledges in other places. Understanding why this is so would help us shed further light on why a meaning theory should be construed as an account of a speaker's understanding of his own language. Consider the following passage from 'Indeterminism and Antirealism':

First person interpretations are necessarily tied to the homophonic translation manual (which is to say, translation, or interpretation, has no place here) . . . It should not be concluded from the fact that a person is restricted to a unique way of interpreting himself (if this can be called interpretation: it would be better to say that aside from pathological cases, our way of interpreting others has no application to ourselves) that therefore his words have unique reference. (Davidson 2001, 80)

Why exactly does Davidson think that our way of interpreting others has no application to ourselves? We get some clarification in 'First Person Authority,' where Davidson describes the presumption that a speaker knows what his own words mean as 'essential to the nature of interpretation—the process by which we understand the utterances of a speaker.' Davidson continues:

This process cannot be the same for the utterer and for his hearers . . . there can be no general guarantee that a hearer is correctly interpreting a speaker; however easily, automatically, unreflectively, and successfully a hearer understands a speaker, he is liable to serious error. In this special sense, he must always be regarded as interpreting a speaker. The speaker cannot, in the same way, interpret his own words. (Davidson 2001, 12–13)

Why, then, does Davidson think that a speaker cannot be said to interpret his own words in the same way as a hearer can be said to interpret the words of a speaker? This passage suggests that this has something to do with first person authority, or with the general presumption (essential to interpretation) that a speaker knows what he means by the words he uses. By contrast, Davidson wants to describe even the most mundane cases of one person's understanding of another utterance as being

ultimately—even if not explicitly—'based on evidence and inference' (Davidson 2001, 66). Davidson thus seems to want to reserve the term 'interpretation' for those cases that are based on evidence and inference, rather than for the understanding each speaker has of his own utterances.¹⁶

Of course, a speaker might, on occasion, fail to be interpretable, and this is the sense in which he might fail to know what his words mean. However, what Davidson leaves no room for is the idea of a speaker's misunderstanding his own utterance, in the sense of its successfully meaning something while the speaker takes it to mean something else. In 'What is Present to the Mind,' he says:

I can do no better, in stating the truth conditions for my utterance of the sentence 'The Koh-i-noor diamond is a crown jewel' than to say that it is true if and only if the Koh-i-noor diamond is a crown jewel. If I say this, I utter a tautology, but if you give the truth conditions of my utterance using the same words, you are making an empirical claim, though probably a true one. (Davidson 2001, 66)

It is important to understand the sense in which a competent speaker's homophonic attributions of truth conditions to her own sentences are supposed to be 'tautologous.' I do not think that Davidson means to be characterizing such utterances as logical or necessary truths. Rather, his claim is only that they are not empirical, or, not 'based on evidence and inference.' When I

¹⁶Does this mean that Davidson disagrees with Quine's remarks that radical translation begins at home, or with his own previous claim that 'the problem of interpretation is domestic as well as foreign' ('Radical Interpretation,' Davidson 1984, 125)? No, I do not think that there is any conflict here. This is easy to see if we keep in mind the distinction between idiolect and communal language. When Davidson claims that even ordinary communicative situations involve radical interpretation, he is claiming only that the account of radical interpretation can help shed light on the nature of successful communication between what we think of as speakers of 'the same language.' This does not involve thinking of the speaker as being in the position of a radical interpreter with respect to her own utterances.

say that 'Snow is white' as uttered by me now is true if and only if snow is white, the claim I am making is contingently true, but it is one that I know to be true simply by virtue of understanding it. Adding to this Davidson's assumption that when communication is successful, the speaker understands her own utterance (since the utterance means what she takes it to mean), we get the result that as long as a speaker is a successful communicator, she knows the conditions under which her sentences are true. It is the content of this basic knowledge that a meaning theory for a speaker's language should be specifying.

Thus, instead of saying that a translation theory 'leaves tacit and beyond the reach of theory what we need to know that allows us to interpret our own language' (Davidson 1984, 130), Davidson should have said that a translation theory leaves beyond the reach of theory what we know about the truth conditions of our own sentences as we currently understand them.

3.3. Apparent counter-evidence

One might wonder whether the contrast I am drawing between the knowledge involved in speaking and in interpreting a language really comes to anything. For, when communication is successful, both speaker and interpreter can be said to know the conditions under which the speaker's utterances are true. What, then, is the point of my insistence that a meaning theory ought to describe a speaker's, rather than an interpreter's, knowledge of a language? The following passage would indeed appear to suggest that Davidson himself directly opposes any substantive contrast between the two:

Because a speaker necessarily intends first meaning to be grasped by his audience, and it is grasped if communication succeeds, we lose nothing in the investigation of first meaning if we concentrate on the knowledge or ability a hearer must have if he is to interpret a speaker. What the speaker knows must correspond to something the interpreter knows if the speaker is to be understood, since if the

speaker is understood he has been interpreted as he intends to be interpreted. ('A Nice Derangement of Epitaphs,' Davidson 2005b, 92–93)

Does this passage not directly undermine my contrast between a speaker's and an interpreter's knowledge of a language? I do not believe it does, for in contrasting a speaker's knowledge of her own language with the knowledge on which an interpreter's understanding of the speaker might rest, I am not claiming that what a speaker knows about each of her sentences goes beyond what a successful interpreter of the speaker knows about these sentences: among the things that an interpreter could know that would put him in a position to understand a speaker are the facts entailed by an acceptable translation theory from the speaker's language onto his own. If and when the interpreter knows these facts, what he knows, in knowing them, does not correspond to anything the speaker *needs* to know in order to be understood (though the speaker *may* well know these facts). However, when the speaker is understood, she and her interpreter share knowledge of the conditions under which her sentences are true, even when the interpreter's knowledge of these truth conditions is based on something else he knows about the speaker's sentences that the speaker herself does not know.

Thus, my insistence that a theory of meaning for a speaker is primarily a description of the speaker's knowledge of her language—rather than of knowledge that would enable us to interpret the speaker—does not involve any commitment to any features of meaning or understanding that are only available from the first-person perspective. That would be patently at odds with Davidson's view of the social and external dimension of language:

It is the requirement of learnability, interpretability, that provides the irreducible social factor, and that shows why someone can't mean something by his words that can't be correctly deciphered by another. ('Knowing One's Own Mind', Davidson 2001, 28)

[T]he objects to which we relate people in order to describe their attitudes need not in any sense be *psychological* objects, objects to be grasped, known, or entertained by the person whose attitudes are described. (Davidson 2001, 36)

In claiming that a theory of meaning ought to capture the content of a speaker's knowledge of the truth conditions of her sentences, I am not construing this knowledge as involving grasp of any objects before the mind. Facts about what a speaker knows about the conditions under which her sentences are true are no less intersubjectively available than facts about the conditions under which the speaker holds various sentences true, or facts about the external causes of her beliefs. My distinction between a speaker's knowledge of her own language and knowledge that could put others in a position to understand her thus does not involve reinstating any subjective elements in meaning, any more than does Davidson's own insistence on the primacy of the idiolect.

Let me spell out how my proposal bears on the project of radical interpretation. If I am right, what a meaning theory in the guise of a truth theory is meant to capture is a speaker's knowledge of the truth conditions of her sentences—knowledge that she relies on in speaking the way she does. This is perfectly compatible with Davidson's insistence that we can shed light on meaning by asking how a theory could be confirmed as acceptable for a given speaker, on the basis of evidence available in radical interpretation. Only, my proposal does require a particular way of construing the task. If I am right, what we are confirming, in confirming a truth theory as acceptable for a given speaker (i.e., as an empirically adequate description of her language on a particular occasion), is not just the truth of its T-theorems, but the plausibility of attributing to the speaker knowledge of these theorems. For example, if an acceptable truth theory for the speaker entails that the sentence 'Londres est belle' as uttered by the speaker is true if and only if London is beautiful, this is something I can take the speaker to know about her sentence. Conversely, if the available evidence were to make

our attribution of such knowledge implausible, this would count against the theory's suitability as a meaning theory.

Finally, my proposal does not involve denying that a speaker's ability to make meaningful utterances depends on her grasp of an intersubjective notion of truth, and therefore on her being an interpreter of others (see 'The Second Person,' Davidson 2001, 121). Quite the contrary: if the arguments of the previous sections are right, it is only by thinking of a truth theory as an account of the speaker's knowledge of her language that we can explain why this theory needs to employ an intersubjective notion of truth. Of course, if I am right, a meaning theory for a speaker captures her knowledge of the truth conditions of her own sentences, but it does not directly capture either her knowledge of the truth conditions of the sentences of others, or the knowledge that her interpreters rely on (or could rely on) in interpreting her utterances. But given Davidson's idiolectal conception of language, this is as it should be.

4. Conclusion

Davidson clearly intended his account of radical interpretation to illuminate the notion of meaning, along with those of truth, rationality, and the propositional attitudes. He did not, however, seek to ground each and every one of his proposals about these notions in considerations about the ends and means of interpretation. A case in point is his proposal to use truth theories as meaning theories.¹⁷ As I argued in this paper, and contrary to widespread assumptions, the point of a meaning theory, and the sense in which a truth theory can serve as one, cannot be explained in terms of its role in interpretation.

The prevailing assumption that it is to be so explained has had the unfortunate consequence of obfuscating a clear view of the

¹⁷Another example is Davidson's commitment to the compositionality of language. See Davidson (1999).

role of the concept of truth in a Davidsonian approach to meaning. For, as we have seen, this assumption naturally leads to an instrumentalist view of the role of truth in truth-conditional meaning theories, thus undermining the contrast between meaning theories and translation manuals. The main point of a meaning theory for a given speaker's language is not to enable us to interpret utterances in it, but rather, to give us a compositional account of the speaker's knowledge of her own language. Once we shift our attention from interpreter to speaker, we can begin to understand the significance of the notion of truth in a Davidsonian meaning theory, and this brings us closer to understanding Davidson's opposition to deflationary conceptions of truth.

Acknowledgements

I would like to thank two anonymous reviewers and the editors of this journal for useful feedback. I am very grateful to Barry Stroud for invaluable discussion of a number of issues addressed in this paper. I would also like to thank Hannah Ginsborg, John MacFarlane and Seth Yalcin for suggestions that greatly improved earlier versions of the manuscript. A precursor of this paper was delivered at the Triennial Congress of the Société de Philosophie Analytique in June 2015. I am grateful to the audience for discussion.

Arpy Khatchirian
University of California Berkeley
arpy@berkeley.edu

References

- Davidson, Donald, 1984. *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- , 1990. 'The Structure and Content of Truth.' *The Journal of Philosophy* 87: 279–328.
- , 1999. 'Reply to Peter Pagin.' In *Donald Davidson: Truth, Meaning, and Knowledge*, edited by U. Zeglen, pp. 68–70. London: Routledge.
- , 2001. *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.
- , 2005a. *Truth and Predication*. Cambridge, MA: Harvard University Press.
- , 2005b. *Truth, Language and History*. Oxford: Clarendon Press.
- Davies, Martin, 1987. 'Tacit Knowledge and Semantic Theory: Can a Five Per Cent Difference Matter?' *Mind* 96: 441–62.
- Evans, Gareth, 1981. 'Semantic Theory and Tacit Knowledge.' In *Wittgenstein: To Follow a Rule*, edited by S. Holtzman and C. Leich, pp. 118–40. London: Routledge and Kegan Paul.
- Foster, John A., 1976. 'Meaning and Truth Theory.' In *Truth and Meaning: Essays in Semantics*, edited by G. Evans and J. McDowell, pp. 1–32. Oxford: Oxford University Press.
- Heck, Richard, 2006. 'Reason and Language.' In *McDowell and His Critics*, edited by C. Macdonald and G. Macdonald, pp. 22–45. Oxford: Blackwell.
- , 2007. 'Meaning and Truth-Conditions.' In *Truth and Speech Acts: Studies in the Philosophy of Language*, edited by D. Greimann and G. Siegart, pp. 349–76. London: Routledge.

- Higginbotham, James, 1989. 'Knowledge of Reference.' In *Reflections on Chomsky*, edited by A. George, pp. 153–74. Oxford: Blackwell.
- , 1992. 'Truth and Understanding.' *Philosophical Studies* 65: 3–16.
- Kölbel, Max, 2001. 'Two Dogmas of Davidsonian Semantics.' *The Journal of Philosophy* 98: 613–35.
- Larson, Richard and Gabriel Segal, 1995. *Knowledge of Meaning: An Introduction to Semantic Theory*. Cambridge, MA: Bradford Books.
- Lepore, Ernest, 1983. 'What Model Theoretic Semantics Cannot Do?' *Synthese* 54: 167–87.
- , 1997. 'Conditions on Understanding Language.' *Proceedings of the Aristotelian Society* n. s. 97: 41–60.
- Lepore, Ernest and Kirk Ludwig, 2005. *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford: Clarendon Press.
- , 2011. 'Truth and Meaning Redux.' *Philosophical Studies* 154: 251–77.
- McDowell, John, 1998. *Meaning, Knowledge, and Reality*. Cambridge, MA: Harvard University Press.
- Quine, W. V., 1981. *Theories and Things*. Cambridge, MA: Belknap Press.
- Ramberg, Bjørn, 1989. *Donald Davidson's Philosophy of Language: an Introduction*. Oxford: Blackwell.
- Rumfitt, Ian, 1995. 'Truth-Conditions and Communication.' *Mind* 104: 827–62.
- Smith, Barry, 1992. 'Understanding Language.' *Proceedings of the Aristotelian Society* n. s. 92: 109–41.
- Soames, Scott, 2008. 'Truth and Meaning—In Perspective.' *Midwest Studies in Philosophy* 32: 1–19.
- Tarski, Alfred, 1944. 'The Semantic Conception of Truth.' *Philosophy and Phenomenological Research* 4: 341–75.
- , 1956. 'The Concept of Truth in Formalized Languages,' translated by J. H. Woodger. In *Logic, Semantics, and Metamathematics*, pp. 152–278. Oxford: Clarendon Press.
- Williams, Michael, 1999. 'Meaning and Deflationary Truth.' *The Journal of Philosophy* 96: 545–64.