The Concept of Truth

Michael Glanzberg
Northwestern University

For the *Companion to Donald Davidson*

Davidson discussed the concept of truth consistently throughout his career. It is central to the early Davidson (1967), and to the late Davidson (2005a). His views on the subject are not absolutely static. At one point, he endorses a version of the correspondence theory of truth, and he comes to emphatically reject that view later. At one point he calls his view a coherence theory of truth, and he later comes to doubt that characterization. Even so, over the period of nearly 40 years that he wrestled with questions about the nature of truth, Davidson is remarkably consistent. Certain points appear early in his thinking and remain constant: Tarski's work on truth is fundamental to understanding the concept, as is the relation of truth to meaning, and we fail to understand that connection adequately unless we take into account the constraints of radical interpretation. Furthermore, in understanding those connections, we see that truth is not a redundant or trivial concept, but equally it is not a locus for metaphysical commitments. Actually, I shall suggest that a stronger conclusion about these points is in order. Not only are they stable points of Davidson's view; they are his view. By his later writings, it becomes clear that we understand the concept of truth through these features of it, and we should not hope for a more direct definition of the concept.

This essay will review Davidson's main work on truth, more or less in chronological order. It will focus on the ways the connections between truth, meaning, and

interpretation form the core of Davidson's views, and on the relations of his views to traditional theories of truth.

1. Tarski, Truth, and Meaning

Any discussion of Davidson's views on truth must begin with the connection between truth and meaning. I discuss this connection, emphasizing what it tells us about Davidson's views on truth. I shall for the most part leave the complex issues of meaning to other discussions.[1]

Davidson's ideas about truth and meaning start with seminal work of Tarski (e.g. 1935; 1944) on truth. Recall, Tarski is concerned with how to characterize truth for a given interpreted language $\mathcal{L}$.[2] To do this, Tarski first introduces an adequacy condition on theories or definitions of truth. It relies on the T-sentences, which are sentences of the form:

's' is true $\leftrightarrow p$

---

[1]For more on Davidsonian approaches to meaning, see Davies (1981) and Lepore and Ludwig (2005; 2007). Some important discussions include the exchange between Higginbotham (1992) and Soames (1992), as well as Larson and Segal (1995).

[2]Of course, this is a somewhat anachronistic way of putting things, and skips a great deal of detail. For a historically accurate discussion, see Patterson (2012).

where 's' is a sentence of $\mathcal{L}$ and 'p' is a sentence of an appropriate metalanguage. In Tarski's original formulation, it is required that 'p' be a translation of 's', but we will see that for Davidson, this condition cannot be assumed. The condition of adequacy on a theory of truth is embodied in 'Convention T', which requires that for each sentence of $\mathcal{L}$, the corresponding T-sentence is derivable. Satisfying Convention T (with the translation requirement in place) guarantees that the extension of 'is true' contains all and only the true sentences of $\mathcal{L}$. Thus, it is a reasonable (extensional) adequacy condition on theories of truth.

It is tempting to think that the T-sentences comprise a definition of truth, and indeed, for finite languages, Tarski considers that they could be treated as such. For a finite language, simply conjoining all the T-sentences is an adequate definition of truth, by lights of Convention T. But for any infinite language, we cannot do this. Instead, we need to find a recursive way to generate the T-sentences.

Tarski provides just such a way, and moreover, a highly natural way to capture how the truth of sentences depends on the semantic properties of their parts. Tarski's method begins with axioms stating the reference conditions for terms and satisfaction conditions for predicates. For a term 't' we should have something like:

't' refers to t.

For a predicate 'P' we want:

x satisfies 'P' $\leftrightarrow Px$.

In addition, we need recursion axioms telling us how satisfaction for complex sentences—including sentences with quantifiers—is determined from satisfaction for their parts. I shall skip the details.[3] However, there are a couple of technical points that will matter for our discussion of Davidson's views of truth. First, because we have predicates with multiple arguments, satisfaction is defined as a relation to a *sequence* σ of assignments of values to variables. As the theory recursively determines satisfaction, what results is a satisfaction relation for formulas. Of course, for a sentence, with no free variables, satisfaction is inert. Thus, for a sentence '*s*', truth can be defined from satisfaction by quantifying out the sequence: '*s*' is true if and only if there is a sequence σ (equivalently, for all sequences σ) such that σ satisfies '*s*'. The theory which includes the sorts of axioms I have just described shows how truth for complex sentences is determined by satisfaction for their parts. Thus, following the terminology of Halbach (2011), we may call it the *Tarskian compositional truth theory* for a language $\mathcal{L}$.[4]

Famously, Tarski is interested in *definitions* of truth, not theories (hence my ambivalence about definitions and theories in introducing Convention T). He shows, essentially, how for a large class of languages, the compositional truth theory can be converted into an explicit definition of truth for the language. Roughly, using a second-

[3]See Lepore and Ludwig (2007) or Soames (1999) for more thorough presentations.

[4]For another technical review of Tarskian truth theories, see McGee (1991).

order quantifier, one defines the truth predicate to be the smallest collection making the compositional truth theory come out true.[5]

With this background on theories of truth in hand, we can turn to Davidson's use of them. In a widely discussed series of papers (Davidson 1966; 1967; 1970; 1973a; 1976), Davidson makes the proposal that a Tarskian compositional truth theory plays a central role in a theory of meaning. To simplify somewhat, a Tarskian compositional truth theory *is* a theory of meaning.

Davidson relies on two aspects of the Tarskian apparatus to create the link between truth and meaning: the T-sentences and the compositional theory. First, a correct T-sentence, like ' 'snow is white' is true if and only if snow is white', does more than indicate that the sentence is true; it states the truth *conditions* of the sentence, and moreover, when the sentence on the right-hand-side is really the right one, it goes a long way towards stating the meaning of the sentence on the left-hand-side. Davidson is careful about how he puts this, in part because he holds rather holistic views about meaning, but he does propose that a theory which provides the right T-sentences will be a good theory of meaning. Thus, a theory of *truth* is the core of a theory of meaning.

There are some departures from Tarski's original formulation that are required to make Davidson's proposal work. As I mentioned above, if we are to use a theory of truth as a theory of meaning, we cannot build translation into the T-sentences. Hence, I have described things in terms of 'the right T-sentences'. Just how to capture which are the right

[5]For more discussion of the role of definitions in Tarski's work, see Heck (1997).

T-sentences has been an issue of ongoing debate for neo-Davidsonians. As we will discuss more below, Davidson himself often shifts the focus to the general question of the empirical constraints on attributing a truth theory to a speaker. Davidson's ideas on radical interpretation are fundamental to how he answers this question, and we will explore how this affects his views on truth in section 3.

The compositional truth theory is also important to Davidson's views. Tarski shows how, for a language with infinitely many sentences, we can derive all the T-sentences from such a theory. In Davidson's hands, this shows us several things about meaning. It shows how the meanings of sentences depend on the meanings of their parts (Davidson 1967, p. 23), and that in turn shows how an infinite language could be learnable by a creature like us with a finite mind (Davidson 1966). This is in part why Davidson is so optimistic about the use of Tarskian truth theories, concluding (Davidson 1967, p. 24), "a Tarski-type truth definition supplies all we have asked so far of a theory of meaning."

One issue that has come up repeatedly in the literature on Davidson is whether this equation of a truth *definition* with a theory of meaning embodies a mistake. The worry is that Tarski shows how to define a truth predicate as an exercise in logic or mathematics for interpreted languages. But then, the T-sentences that Davidson sees as showing that we have a plausible theory of meaning are really facts of pure mathematics or logic. How, it is objected, can a collection of facts of pure mathematics or logic be a good theory of meaning

for a language like the ones we speak? Surely, the facts about what our words and sentences mean are empirical, not mathematical.[6]

Davidson has discussed this point in several places, and at length in Davidson (1990). It is arguable that in Davidson (1967) he might have indeed made the mistake he is accused of, and the language quoted above about Tarskian truth *definitions* does make that reading plausible. But by the time he wrote the introduction to Davidson (2001a), he was clear that what plays a role in a theory of meaning is a truth *theory*, not an explicit definition of truth. Though Tarski was interested in the latter, Davidson is only interested in the former. When he talks about a Tarskian theory of truth, he means what I called the compositional truth theory, not the explicit definition we might build from it. In such a theory, the term 'true' is a primitive. But moreover, as is typically the case with primitives of empirical theories, we assume they are given in an interpreted metalanguage, and so the term 'true' in the compositional truth theory really means true. Then, the axioms of the theory can perfectly well state empirical facts about the truth conditions of sentences in the object language. I thus believe that Davidson should be cleared of the charge of confusing the mathematical with the empirical, at least in his later writings.

We now have one of Davidson's fundamental theses about truth: a Tarskian compositional truth theory plays a key role in a theory of meaning. So far, we have one thesis from Davidson connecting truth and meaning. As things looked in 1967, though, it

---

[6]This sort of objection has been voiced often. In the recent literature, it can be seen in Etchemendy (1988) and Soames (1984). See also Putnam (1985–86).

seemed that Davidson had little interest in exploring the concept of truth itself. The correct understanding of Davidson's thesis linking truth and meaning requires *using* the concept of truth in an empirical theory of meaning, not analyzing it. Ultimately, as we will see in the following sections, Davidson is interested in the nature of truth itself, and he sees the link between truth and meaning as a key step towards understanding that nature.

2. Satisfaction and Correspondence

One way to get a more substantial theory of truth out of the link between truth and meaning is to see a Tarskian compositional theory of truth as providing one. In particular, the Tarskian theory has sometimes been seen as giving a correspondence theory of truth. Early on, Davidson endorsed this proposal. It is not his considered view on the matter, but it is an important episode in the Development of Davidson's views. In this section, we will review the episode.

One of the leading ideas about truth over many years has been the 'correspondence theory of truth'. Very roughly, and informally, such a theory proposes that for a truth bearer to be true is for it to 'correspond to' or correctly describe what really is the case. There are two ways one can view this idea. It can be taken as something near to a platitude about truth, and many have taken it as such (e.g. Wright 1992). But it is also the beginnings of a metaphysically substantial theory of truth: one which seeks to answer directly the question of what the basic nature of truth is. The platitude already posits that truth is a relation, between something appropriate for bearing truth—something that is or has a content, like a statement, proposition, etc.—and some extra-intentional part of reality to

which the representation corresponds. The traditional correspondence theory, as I shall call it, seeks to develop this idea into a full-fledged metaphysical theory of the nature of truth.

Tarski's work on truth has been seen by commentators as supporting a number of different views on truth, including correspondence theories. Tarski himself endorses at least the platitude behind the correspondence theory, saying (Tarski 1935, p. 153) that his work on truth captures the "classical conception of truth ('true—corresponding with reality')." T-sentences, as Davidson and others have noted, express a relation between a sentence and something extra-linguistic, which would seem to capture in precise terms what the platitude might be saying. But moreover, many readers of Tarski have thought that the role of satisfaction in the compositional theory of truth makes for a more substantial sort of correspondence theory.

Davidson (1969) takes up this idea and endorses it (though he later comes to change his mind). In that paper, he argues that a Tarskian composition theory is a non-traditional but substantial kind of correspondence theory of truth. Let us begin with the non-traditional aspects. Davidson rejects key components of the traditional correspondence theory, and indeed, the negative aspects of his position in 1969 endure through his later work. The traditional correspondence theory, in the hands of Russell (1912) or Moore (1953), posited *facts* as the things to which a truth-bearer corresponds. Facts are structured entities, which are in effect small pieces of reality suitable for being truth-makers for specific claims. The traditional correspondence theory is based on the idea of a structural similarity relation (an isomorphism) between facts and suitable truth

bearers. The relation of structural similarity is the relation of correspondence. When Russell and Moore held the correspondence theory, those truth bearers were beliefs, though many have offered propositions, or sentences, as the truth bearers for the correspondence theory. Following the terminology of Davidson (1969), we can simply talk about 'the statement that *s*' as a truth bearer, without worrying too much about what a statement is. So, by Davidson's lights, a key component of the traditional correspondence theory is the idea that for a given sentence '*s*', there is an appropriate fact, let us say the fact that *p*, to which the statement that *s* corresponds if it is true.

Davidson rejects any such role for facts. In particular, he rejects the idea that we can make sense of a particular fact corresponding to a statement. In a number of places, including Davidson (1967; 1969; 1990), he cites an argument which seeks to show this, by showing that if a true statement corresponds to a fact, it corresponds to the fact reported by any true sentence. For any true sentences '*s*' and '*t*', the argument claims, the statement that *s* corresponds to the fact that *t*. If this holds, the idea of truth as a relation between statements (truth bearers) and the facts they characterize loses its substance, and the traditional correspondence theory is badly damaged.

This sort of argument has a long history. Davidson attributes it to Church (1943; 1956), but also notes its connection to ideas of Frege (1892). Related arguments can be found in Quine (1953) and importantly Gödel (1944). Here is Davidson's (1969, p.42) version.[7] The claim, as I said, is that for any true sentences '*s*' and '*t*', the statement that *s*

[7]The argument was dubbed the 'slingshot' by Barwise and Perry (1981)

corresponds to the fact that t. Thus, any true statement corresponds to any fact. The argument in Davidson's hands relies on two premises: (1) if 'p' and 'q' are logically equivalent, then the statement that s corresponds to the fact that p if and only if the statement that s corresponds to the fact that q, and (2) if 'p' and 'q' differ only by substitution of coreferential singular terms, then the statement that s corresponds to the fact that p if and only if the statement that s corresponds to the fact that q. Now, let 's' be some true sentence. It is taken for granted that the statement that s corresponds to the fact that s. Observe that 'the x such that (x = Diogenes and s) = the x such that (x = Diogenes)' is logically equivalent to 's'. Call this complicated sentence 'L'. By premise (1), we can conclude that the statement that s corresponds to the fact that L. Suppose 't' is true. Then 'the x such that (x = Diogenes and s)' is coreferential with 'the x such that (x = Diogenes and t)'. Let 'L'' be 'the x such that (x = Diogenes and t) = the x such that (x = Diogenes)'. Then 'L' and 'L'' differ only in the substitution of coreferential terms.[8] Hence, by premise (2), the statement that s corresponds to the fact that L'. Finally, note that 'L'' is logically equivalent to 't', so by one more application of premise (1), we conclude that the statement that s corresponds to the fact that t.

We thus have an argument for the collapse of the notion of fact: any fact is as good as any other for standing in the correspondence relation to sentences. We might suppose there is one 'great fact', which conjoins all the facts, and any true sentence corresponds to

[8]Issues about the nature of descriptions are crucial here, as Neale (2001) discusses at length.

that. Of course, there are a number of places where the premises of the argument might be rejected,[9] but, to Davidson, it provides good reason to reject the apparatus of facts that was at the core of the traditional correspondence theory.

Davidson thus rejects one of the key tenets of the traditional correspondence theory, but even so, in 1969 he endorses a non-traditional version which he thinks still deserves the name 'correspondence theory'. Not surprisingly, it is in Tarski's theory of truth that he sees the possibility of a correspondence theory. Indeed, he says that the compositional theory "deserves to be called a correspondence theory" (Davidson 1969, p. 48) because of the role played by satisfaction. A Tarskian compositional truth theory determines truth for sentences via reference and satisfaction. These are relations "between language and something else" (Davidson 1969, p. 48), and one central aspect of the correspondence theory was that it tried to capture such relations. Though we do not have a traditional correspondence theory, which sets up a structural similarity relation between statements and facts, we have the truth of sentences determined by word-to-world relations; in particular, relations between linguistic items and objects. This reflects the platitude behind correspondence, but the Tarskian apparatus gives the platitude more substance.

This provides a way to see truth as determined by satisfaction, but it seems to leave out any relation between sentences and anything else. Even if having relations of satisfaction constitutes a form of correspondence, it is thus drastically different from the

---

[9]See Neale (2001) for an extensive discussion. Some responses from theorists of facts (or related notions) include Barwise and Perry (1981) and Vision (2004).

traditional form which posited truth-making relata for each statement. But Davidson is happy to extend the sort of correspondence he finds to whole sentences (without free variables). As we discussed above, we still have the relation of σ satisfying '*s*' for a sentence '*s*' and sequence σ. This, Davidson suggests, is a correspondence relation for sentences, just as it is for terms. A sentence is true if it is satisfied by any or all sequences. If we think of satisfaction as correspondence, a sentence is true if it corresponds to any or all sequences. We have built in the 'great fact', but for sequences of objects, rather than facts per se. According to Davidson, we thereby have all the correspondence we should want.

So, at least at one point, Davidson does offer a form of a correspondence theory. Whether or not it is much like its traditional counterpart, it is striking that Davidson offers a theory that addresses the same questions as a traditional metaphysical theories of truth. All the same, Davidson comes to reject the correspondence theory. To see why, and to see what replaces it, we need to consider another aspect of his views about meaning: the important role for radical interpretation.

3. Radical Interpretation and Coherence

Davidson's ideas about radical interpretation and its place in the theory of meaning are fundamental to the way his view of truth develops. I shall not here attempt to review all the important aspects of Davidson's notion of radical interpretation and its uses. But, I shall outline enough of them to examine the impact of radical interpretation on his understanding of truth.

In early papers like Davidson (1967), he highlights the role of a truth theory in the theory of meaning. But as his thinking developed, Davidson came to place more emphasis on the question of how a theory of meaning—now in the form of a truth theory—might be correctly attributed to a speaker. Knowing that a theory of truth can play the role of a theory of meaning does not by itself tell you what truth theory captures a speaker's meanings, or how you should decide what truth theory to use to interpret a speaker.

In addressing these issues, of course, Davidson is heavily and explicitly influenced by Quine (e.g. 1960). Following Quine, he puts great emphasis on the kinds of publicly available evidence that an interpreter might use to understand a speaker (though he explicitly distances himself from Quine's behaviorist take on this idea). To see what that evidence can support, we consider the situation of radical interpretation, where an interpreter knows nothing about what a speaker believes or thinks, and tries to interpret them.

The basic evidence that Davidson sees as available for radical interpretation is the attitude of *holding true*, which a speaker takes towards a sentence (cf. Quine's use of assent and dissent behavior). This evidence can be used to assess whether T-sentences are correct for a speaker, i.e. what the truth conditions of (some) sentences used by the speaker are (e.g. Davidson 1973b; 1974). Importantly, this sort of evidence applies at the level of sentences and their truth conditions, i.e. at the level of Convention T. It does not apply at the level of individual facts about reference or satisfaction that are used to build up a Tarskian compositional truth theory. But this is all the evidence there is, and so, Davidson concludes, the compositional apparatus outstrips the evidence. He writes, (1973b, p.133),

"All this apparatus is properly viewed as theoretical construction, beyond the reach of direct verification. It has done its work provided only it entails testable results in the form of T-sentences, and these make no mention of the machinery." (He repeats this point in Davidson 1977.)

As Davidson himself is well aware, this conclusion undermine the position he took in Davidson (1969) that Tarskian compositional truth theories are correspondence theories of truth. That position relied on the use of reference and satisfaction to provide word-to-world relations. But now, these relations are seen as nothing but theoretical constructs with no empirical content. They outstrip the possible evidence, and so, do not report any real facts of the matter. Thus, there is nothing left of the correspondence theory except such an empirically empty theoretical construct. At one point Davidson (1986), he suggests this means we have a sort of 'lightweight' correspondence theory, but his more considered view is that we should reject talk of correspondence. He explicitly says so in the "Afterthoughts" to Davidson (1986), he calls his earlier view a "mistake" (Davidson 1988), and expresses "regret" (Davidson 1990). At one point (Davidson 1990), he notes the "oddity" of holding that sequences of objects stand in a correspondence relation to sentences. That seems to be something of a technical trick, and not a real version of correspondence. But the reasons Davidson rejects the correspondence theory go much further than that, as he sees there being no facts of the matter about satisfaction at the subsentential level.

Obviously, Davidson's argument only goes through if we accept his view that the facts about meaning are fundamentally exhausted by the kinds of information available to

radical interpretation, especially, facts like those about holding-true. A number of other views about meaning, including some in the neo-Davidsonian tradition, will not accept this.[10] When it comes to the theory of truth, an alternative view was defended by Field (1972). Field, argues that a Tarskian theory of truth needs to be supplemented with a reductive account of reference and satisfaction. If we had such an account, then the kinds of word-to-world relations Davidson originally relied upon would be substantiated, and we would have a correspondence theory in much the form Davidson originally envisaged.

The approach taken in Field (1972) calls for a reductive account of reference, and he argues that without such an account, a Tarskian theory loses its value. But, one might imagine other positions as well. We might, for instance, imagine a more direct sort of response in the manner of scientific realism, which might insist that insofar as a Tarskian theory is confirmed, so too are all the facts stated by the apparatus it uses. We might also ask if Davidson is really right about what the possible sources of evidence might be. Whatever the other possibilities may be, Davidson does not accept them, and the only substance Davidson comes to see for the compositional apparatus of reference and satisfaction is its instrumental role in building theories which can derive sufficiently many T-sentences.

The basic constraints of radical interpretation lead Davidson to reject the idea that the role of a composition truth theory in the theory of meaning could substantiate a correspondence view. But this does not lead Davidson to give up on trying to make sense of

---

[10]See, for instance, Higginbotham (1989).

traditional theories of truth within the broader setting of relating truth and meaning. The role of radical interpretation points Davidson in a very different direction than correspondence theories. At one point, he entertains what he calls a coherence theory of truth on the basis of considerations of radical interpretation. He ultimately comes to reject this proposal too, but it comes closer to his considered position. We will discuss these points for the remainder of this section.

Whenever he discusses radical interpretation, Davidson is quick to point out that it must simultaneously identify what a speaker believes and what they mean. When we assume we have evidence that a speaker holds a sentence true, it is simultaneously evidence of what the speaker means by that sentence, and what the speaker thinks: the sentences means something that the speaker thinks is true in the circumstances. But to produce a theory of meaning, we need to disentangle these two factors. We do so by interpreting the speaker as broadly holding true beliefs, judged, of course, from the perspective of the interpreter (Davidson 1973b; 1974). Thus, we interpret the speaker as holding beliefs that broadly match up with our own. This affects radical interpretation in at least two ways. First, when we come to asses an individual case of holding true, we rely on our own beliefs (which we take to be true!) to factor belief from meaning. If I believe it is raining near Kurt and hear Kurt utter 'Es Regnet', it is evidence that 'Es regnet' is true if and only if it is raining (near the speaker, etc.). But also, the instruction to interpret speakers as holding true beliefs functions as a maximizing principle for interpretation (as highlighted in Davidson 1973b). As a strategy for radical interpretation, we should interpret speakers so as to have their beliefs come out true as much as is possible. This is a *principle of charity* for radical interpretation.

The principle of charity can be seen as simply a strategy for solving the problems of radical interpretation. But in Davidson's hands, it becomes more. The reason is that there are, to Davidson, no other facts beyond those revealed by radical interpretation which could fix meaning or belief. It is radical interpretation, and only radical interpretation, which reveals the nature of belief and meaning, and sustains the connection between truth and meaning. There is thus no other check on truth than what we get from the process of radical interpretation. We thus conclude not only that our strategy for radical interpretation is to interpret speakers as broadly holding beliefs which are true by our lights, but that they really do hold beliefs which are broadly true by our lights. This is part of the nature of belief. Furthermore, the proviso 'by our lights' does not really matter. To make this point vivid, Davidson (1986) asks us to suppose there was some other interpreter than us, who was omniscient. Such an interpreter would find the beliefs of any subject of interpretation mostly true, as we would. Indeed, the omniscient interpreter would find our beliefs mostly true. Hence, our beliefs are mostly true, by the very nature of belief, and so are the beliefs of those we interpret. But we do not really need to appeal to omniscience to get this result. The nature of belief—ours, those we interpret, those who interpret us, omniscient or not—is revealed by radical interpretation, and that builds in the veridicality of belief. The veridicality of belief implies the surprising and much discussed thesis that most of the beliefs held by a speaker are true. This thesis is a key lemma for many of Davidson's subsequent arguments. So, I shall call it the *Davidsonian lemma*.

The Davidsonian lemma, as we will see, invites the view that truth is epistemic in nature. The traditional version of this view is the coherence theory of truth. At one point, Davidson (1986) entertains the idea that a coherence theory of truth (at least, a form of

one) could be defended on the basis of the Davidsonian lemma. He soon modifies this position, and in later work considers and substantially rejects other epistemic views of truth, as we will discuss more in section 5. Yet, even if it does not support a coherence theory, the lemma provides another fundamental insight into truth, along with the association of truth and meaning.

The coherence theory of truth is another traditional theory, and indeed was often the main competitor to correspondence theories. Very roughly, the traditional coherence theory holds that to be true is to be part of a coherent set (of appropriate truth bearers). This view is closely associated with forms of idealism. Like many long-standing philosophical ideas, genuine forms of the coherence theory were far more sophisticated than the slogan that truth is coherence indicates.[11] However, this overly simple slogan will be enough to discuss Davidson's own views, and to see why the coherence theory is a genuine alternative to the correspondence theory.

The connection between the Davidsonian lemma and the coherence theory of truth is tantalizing, but also rather delicate. The lemma shows something about belief (and meaning): it is, as Davidson says, "intrinsically veridical" (Davidson 1986). But the beliefs in question are really systems of belief, as radical interpretation interprets the whole language of a speaker, and gets to the veridical nature of belief by maximizing true belief

---

[11]For instance, the much discussed view of Joachim (1906) has a complex notion of 'coherence' relying not only on idealism but a form of monism. For an extended discussion of coherence theories, see Walker (1989).

across all the agents beliefs and meanings. So, it would appear that beliefs become veridical in nature by being part of the systems of belief held by agents. This does not quite get us the traditional coherence theory. In this setting, a version of the traditional theory would tell us that any one belief is true in virtue of being part of a system of beliefs under radical interpretation, which would render it part of a coherent and otherwise rational system. What we get from Davidson is less than this. We only know that each belief is veridical in nature, in virtue of being part of such a coherent system. We cannot conclude that each belief individually is true. Only the weaker conclusion that most of a speaker's beliefs are true follows. That is in the spirit of a coherence theory, but not as strong as the traditional version.

In Davidson (1986), Davidson still thought the result was worth calling a coherence theory of truth and of knowledge. The reason was that even if not the traditional coherence theory, the connection between veridicality and coherence is a substantial insight into the nature of truth. Moreover, this connection gives truth significant epistemological properties, as the traditional coherence theory did. Belief enjoys a kind of justification because of its veridical nature. According to Davidson, any sufficiently sophisticated thinker who asks whether their beliefs are justified may understand the nature of belief, and so conclude they are mostly true in virtue of what belief is. The conclusion that most beliefs are true thus requires no further justification. Of course individual beliefs might require some additional justification, but the general conclusion still holds. This is enough to draw some significant epistemological conclusions. For instance, it is enough to reject global skepticism. Thus, the coherence-like aspects of radical interpretation show us

something important about truth and justification. Truth thereby acquires at least some epistemic properties.

Not long after he wrote Davidson (1986), Davidson decided it was better not to call the view a coherence theory of truth. As he says in the "Afterthoughts" to that paper, his position is not really a coherence theory of truth of the traditional kind. As we already saw, it does not say that being part of a coherent set of beliefs makes any one belief true. But more importantly, Davidson's view does not provide any direct answer to the question of what truth is. That was what the traditional coherence theory was trying to do, and in not doing it, Davidson does not offer a theory which should be called a coherence theory. Yet as Davidson's view of truth develops, the veridical nature of belief, and the role of radical interpretation in understanding truth, remain fundamental. As this is a reflection of an idea of coherence, we can say that aspects of coherence are an important part of Davidson's view, even if he does not hold a traditional coherence theory.

We can now list two basic theses of Davidson on truth: truth is fundamentally connected to meaning, and the facts about how truth connects to meaning are limited by what is available in radical interpretation. These later facts give empirical substance to attributing Tarskian compositional truth theories to speakers, and it there that we find out what we can about the nature of truth. We have also seen that Davidson in two separate instances considered identifying his own views of truth with traditional metaphysical theories of truth. In each case, he eventually rejected the identification, but all the same we see how Davidson is concerned with traditional questions in his thinking about truth. In the

next section, I shall turn to how far Davidson thinks the traditional questions really can be answered.

4. Defining Truth and Deflationism

In each case where Davidson considers allying his own views of truth with traditional theories, he eventually decides the facts needed to do so are not to be found. In later work, Davidson comes to conclude that we cannot define truth, and so, the aims of the traditional theories cannot be met. Indeed, he describes the project of defining truth as a "folly." But at the same time, Davidson rejects contemporary deflationist positions which conclude that truth has no nature. To better understand Davidson's later views on truth, we should examine how, and why, he holds both positions.

First, let us consider Davidson's rejection of deflationism. His objections to this family of views are of a very general variety. He takes the main idea of deflationism to be that Tarski has told us all there is to know about truth, and perhaps more explicitly, that the T-sentences tell us all there is to know about truth. Along with this, he notes, goes the idea that truth is not a substantial concept, and does not have an "underlying essence" (Horwich 1990, p. 2) or is not "deep or interesting" (Davidson 2005a, p. 18).[12]

---

[12]At this level of generality, Davidson's objections to deflationism cover a wide range of more specific views. He has discussed some of these in more detail. The early views of Ramsey (1927) and Strawson (1950) are mentioned often in his writings, while Field

Davidson's response to the idea that the T-sentences show us all there is to the concept of truth is quite direct. That cannot be, Davidson observes, because Tarski's work does not show us how to characterize truth in general. It only shows us how to characterize truth for individual languages. As he writes (Davidson 1996, 26), "[Tarski] has not told us what the concept is that his truth definitions for particular languages have in common." Indeed, as Davidson notes, according to one way of taking Tarski's undefinability theorem it shows that no such general definition can be given. Hence, according to Davidson, we should reject any view which sees a general characterization of truth in the T-sentences or other Tarskian apparatus. This includes a number of contemporary forms of deflationism.[13]

What of the more general idea that truth lacks an underlying nature, or other 'essence' that philosophical theories like the correspondence theory or the coherence theory sought to capture? Davidson is generally in favor of stripping the concept of truth of some of its metaphysical weight, which some broadly deflationist view such as those of Soames (1984) emphasized. But he resists the full deflationist idea that truth is a trivial concept. We have already seen why. Truth, to Davidson, is a fundamental concept which interrelates with those of meaning and belief (and perhaps also other basic concepts, like knowledge, cause, or action). So, for instance, he writes (Davidson 2005a, p.28), "The

---

(1994), Horwich (1990), Leeds (1978), Quine (1970), and Williams (1986) are discussed in later work.

[13]In Davidson (1990) he associates with this objection with the famous claim of Dummett (1959) that the T-sentences fail to explain the *point* of classifying something as true.

concept of truth has essential connections with the concepts of belief and meaning, but these connections are untouched by Tarski's work." In these connections, and in the kind of facts that support them, we find substantial properties of truth. Deflationists miss these properties and their source.

But if that is right, then should we not persevere with the traditional quest to find a definition of truth, or at least a theory which directly explains what the nature of truth is? According to Davidson, we should not. We will not, he suggests, do better than isolating the kinds of connections we have seen, and the kinds of facts that support them. We should not expect a more direct theory or an explicit definition. Why not? Davidson puts the reason quite succinctly (Davidson 1996, p. 20):

> For the most part, the concepts philosophers single out for attention, like truth, knowledge, belief, action, cause, the good and the right, are the most elementary concepts we have, concepts without which (I am inclined to say) we would not have concepts at all. Why then should we expect to be able to reduce these concepts definitionally to other concepts which are simpler, clearer, and more basic?

There is a nature to truth, as Davidson sees it, and we can investigate it; hence deflationism is to be rejected. But we cannot investigate a concept like truth directly, by looking for a definition, or a theory which directly reveals its nature. Rather, we find its nature in the connections between concepts. Thus, Davidson's theses that truth connects to meaning, and that attributions of truth theories are constrained by radical interpretation, really do stand as examples how he thinks we should explore the concept of truth.

There is one other way Davidson thinks we can learn about the concept of truth. We can see its structure in the theories of meaning we develop for specific languages. Tarskian compositional truth theories are the prime example. Davidson writes (2005a, p. 28), "What Tarski has done for us is show in detail how to describe the kind of pattern truth must make, whether in language or in thought." This is another indirect source of knowledge about the concept of truth in general. It is indirect, as each Tarskian theory is a theory for an individual language, not truth in general. Yet from such theories, and from the facts about how they are correctly attributed to speakers, we can sometimes extract general patterns which reveal something of the nature of truth.

5. Objectivity and Truth

We have seen the importance to Davidson of the connections between truth and meaning, and the constraints on meaning flowing from radical interpretation. At some points earlier on, Davidson saw those connections and contraints as indicating traditional theories of truth, including sometimes the correspondence theory and sometimes the coherence theory. But his views about truth definitions and deflationism show that over time, these connections became not the basis for a further theory of truth, but the theory itself. As we saw above, he sees basic concepts like truth as accessible only through the sort of indirect theorizing those connections illustrate.

The traditional theories of truth were theories in metaphysics, and they went hand-in-hand with broader metaphysical views. Coherence theories were associated with idealism, and the correspondence theory with realism. Though he is no friend of

deflationism, Davidson is happy to distance the concept of truth from these sorts of metaphysical commitments. I will conclude, in this section, by reviewing the anti-metaphysical side of Davidson's view of truth.

The background to the question of what the metaphysical content of truth might be was set by the discussion of truth running through the 1970s and 80s. In that discussion, the ideas of correspondence and coherence were generalized into positions we can reasonably label realism and anti-realism. Anti-realist positions endorse some epistemically conditioned notion of truth, such as warranted assertibility or verifiability (e.g. Dummett 1976; Putnam 1981; Wright 1976). Like coherence, these notions are epistemically tractable, and within the powers of human agents to assess (perhaps idealized agents). Just as coherence theories went hand-in-hand with idealism, views of truth like these are identified with more general forms of anti-realism. We can, for purposes of discussing Davidson, group them as 'epistemic' views of truth, as he does.

Realism can be described as the view that the facts we describe are independent of us and our thoughts (cf. Wright 1992). All the epistemic views of truth in some way depart from this idea, and hence get their label of 'anti-realism'. Traditionally, realism has been associated with the correspondence theory of truth, as the appeal to facts might suggest. The idea is that our claims are true or false independently of us, as there are facts about what in the world our words pick out, facts about what those things are like, and those in turn determine whether our claims are true. The correspondence theory in its traditional form tries to capture this view of truth, but it was later observed that the realist implications of the correspondence theory are largely independent of the details of the

traditional theory. So, for instance, Dummett (1976) suggests that the mark of realism is bivalence: epistemic notions of truth do not support bivalence, whereas realist ones do. Though this hardly scratches the surface of the many complicated issues relating realism and truth, let us turn to how Davidson sees these matters.

We know already that Davidson came to reject the correspondence theory. Later, he came to see this as amounting to a general rejection of realist positions. His reason is simply that he does not see any substance to this sort of position without a correspondence theory of truth. As we saw, according to Davidson the correspondence theory fails because there is nothing for sentences to correspond to, and no substantial facts about reference or satisfaction. But without these, Davidson says, the idea of realism is just an empty slogan. As he sees it, a realist position would require there to be things in the world that make our sentences true, and substantial facts about how they do so, and he has rejected both. He writes (Davidson 2005a, p. 42), "The only evident positive sense we can make of this phrase [the real and the true are independent of our beliefs], the only use that consorts with the interpretations of those who prize it, derives fromt he idea of correspondence, and this is an idea without content." We can still maintain such platitudes as that believing something does not make it true, but to Davidson there is no substance to a realist position in the metaphysics of truth.

Does this make Davidson an anti-realist? Anti-realism is marked by accepting some sort of epistemic notion of truth, and ultimately, Davidson concludes that epistemic views of truth are generally unacceptable. He is no anti-realist either. This might seem surprising. Even though Davidson came to reject the traditional coherence theory, his considered

views might seem to drive him towards an epistemic view of truth. As he essentially argued in Davidson (1986), the ultimate source of facts about truth is the evidence speakers can rely on to interpret each-other, as captured by radical interpretation. This sort of evidence is designed to be epistemically accessible to interpreters (which leads to substantial epistemic consequences). Is the result not an epistemic constraint on truth, of just the sort epistemic theories of truth require?

Davidson clearly says no. Epistemic views are false. He does grant that they are merely false, and not unintelligible as he found realist views to be (Davidson 2005a, p. 47). Views that make truth radically epistemic, depending on individual agents ability to verify or find evidence for asserting, make truth far too subjective. As he notes, it might allow that a truth could be "lost" as a person loses some epistemic position. Likewise, it deprives truth of its central role as an intersubjective standard. These, he holds are sufficiently fundamental features of truth that we cannot accept any view which is incompatible with them (e.g. Davidson 2000; Davidson 2005a).

Davidson's position here is subtle and difficult. There are basically epistemic constraints on truth from radical interpretation, but we are not to take the step from those to making truth itself an epistemic notion. To do so would get the basic facts about the concept of truth wrong. Davidson thinks we need not take the step, since he sees the connection between truth and belief created by radical interpretation as a basic one, but not a reduction. This connection is not enough to make truth behave like the epistemic notions of warranted assertibility or verification. Hence, the epistemic implications of

radical interpretation do not make truth thoroughly epistemic, even if they have epistemic consequences for the veridicality of belief.[14]

Davidson rejects both realist and anti-realist positions about truth, and is generally skeptical of there being substantial metaphysical content to the concept of truth. This skeptical view has led some to associate Davidson with the pragmatist tradition. Rorty (1986) has suggested as much. Though Davidson is not completely unhappy with this characterization, he does resist some of its consequences. Especially, the kinds of pragmatist views that make truth the end result of scientific investigation, or what promotes certain kinds of successful outcomes, or is useful, are no more congenial to his views than the traditional metaphysical theories are. They require a normative role for truth which he does not find, and fail to recognize the fundamental relations between truth and meaning which are so central to Davidson's thinking (e.g. Davidson 2000; 2005a).

Davidson takes a different approach to truth than many in the traditional and current debates. Against deflationists, he holds that truth is a substantial concept. Against many traditional metaphysical theories, he holds that we are not going to be able to define or directly analyze truth, while against many contemporary metaphysical views, he holds

---

[14]Here I essentially follow an editorial footnote to Davidson (2005a, p. 75) by Charles Parsons. The delicate nature of Davidson's position is revealed by he himself suggesting in a note to the manuscript of Davidson (2005a) that some form of epistemic view must be right after all. Parsons' note attempts to make sense of this in light of Davidson's own clear rejection of epistemic views earlier in the book.

that what more there is to truth than deflationists admit will not be found in the metaphysics of realism, objectivity, or epistemic notions. He rejects most of the standard positions on truth—both traditional and contemporary—in favor of the view that truth is a basic concept, which we explore through its connections with other concepts. According to Davidson, it is connections between truth, meaning, and belief, seen through the lens of radical interpretation, which give us our understanding of truth.

References

Barwise, J. and J. Perry, 1981. "Semantic innocence and uncompromising situations." In P. A. French, T. E. Uehling Jr., and H. K. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. 6, pp. 387–404. Minneapolis: University of Minnesota Press.

Church, A., 1943. "Review of Quine's 'Notes on existence and necessity'." *Journal of Symbolic Logic* 8: 45–47.

———, 1956. *Introduction to Mathematical Logic*. Princeton: Princeton University Press.

Davidson, D., 1966. "Theories of meaning and learnable languages." In Y. Bar-Hillel (ed.), *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, pp. 383–394. Amsterdam: North-Holland. Reprinted in Davidson (2001a).

———, 1967. "Truth and meaning." *Synthese* 17: 304–323. Reprinted in Davidson (2001a).

———, 1969. "True to the facts." *Journal of Philosophy* 66: 748–764. Reprinted in Davidson (2001a).

———, 1970. "Semantics for natural language." In *Linguaggi nella Società e nella Tecnica*, pp. 177–188. Milan: Comunità. Reprinted in Davidson (2001a).

———, 1973a. “In defense of Convention T.” In H. Leblanc (ed.), *Truth, Syntax and Modality*, pp. 76–86. Amsterdam: North-Holland. Reprinted in Davidson (2001a).

———, 1973b. “Radical interpretation.” *Dialectica* 27: 313–328. Reprinted in Davidson (2001a).

———, 1974. “Belief and the basis of meaning.” *Synthese* 27: 309–323. Reprinted in Davidson (2001a).

———, 1976. “Reply to Foster.” In G. Evans and J. McDowell (eds.), *Truth and Meaning*, pp. 33–41. Oxford: Oxford University Press. Reprinted in Davidson (2001a).

———, 1977. “Reality without reference.” *Dialectica* 31: 247–253. Reprinted in Davidson (2001a).

———, 1986. “A coherence theory of truth and knowledge.” In E. Lepore (ed.), *Truth and Interpretation*, pp. 307–319. Oxford: Basil Blackwell. Reprinted in Davidson (2001b).

———, 1988. “Epistemology and truth.” In *Proceedings of he Fourth Panamerican Philosophy Conference*. National University of Cordoba. Reprinted in Davidson (2001b).

———, 1990. “The structure and content of truth.” *Journal of Philosophy* 87: 279–328. Reprinted in revised form in Davidson (2005a).

———, 1996. “The folly of trying to define truth.” *Journal of Philosophy* 93: 263–278. Reprinted in Davidson (2005b).

———, 2000. “Truth rehabilitated.” In R. B. Brandom (ed.), *Rorty and his Critics*, pp. 65–74. Oxford: Blackwell. Reprinted in Davidson (2005b).

———, 2001a. *Inquiries into Truth and Interpretation*. 2nd edition. Oxford: Oxford University Press. 1st edition 1984.

———, 2001b. *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.

———, 2005a. *Truth and Predication*. Cambridge: Harvard University Press.

———, 2005b. *Truth, Language, and History*. Oxford: Oxford University Press.

Davies, M., 1981. *Meaning, Quantification, Necessity*. London: Routledge and Kegan Paul.

Dummett, M., 1959. "Truth." *Proceedings of the Aristotelian Society* 59: 141–162. Reprinted in Dummett (1978).

———,1976. "What is a theory of meaning? (II)." In G. Evans and J. McDowell (eds.), *Truth and Meaning*, pp. 67–137. Oxford: Oxford University Press. Reprinted in Dummett (1993).

———,1978. *Truth and Other Enigmas*. Cambridge: Harvard University Press.

———,1993. *The Seas of Language*. Oxford: Oxford University Press.

Etchemendy, J., 1988. "Tarski on truth and logical consequence." *Journal of Philosophical Logic* 43: 51–79.

Feferman, S., J. W. Dawson Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, and J. van Heijenoort (eds.), 1990. *Kurt Gödel: Collected Works*, volume II. Oxford: Oxford University Press.

Field, H., 1972. "Tarski's theory of truth." *Journal of Philosophy* 69: 347–375.

———, 1994. “Deflationist views of meaning and content.” *Mind* 103: 249–285.

Frege, G., 1892. “Über Sinn und Bedeutung.” *Zeitschrift für Philosophie und philosophische Kritik* 100: 25–50. References are to the translation as “On sense and meaning” by M. Black, reprinted in Frege (1984).

———, 1984. *Collected Papers on Mathematics, Logic, and Philosophy*. Oxford: Basil Blackwell. Edited by B. McGuiness.

Gödel, K., 1944. “Russell’s mathematical logic.” In P. A. Schilpp (ed.), *The Philosophy of Bertrand Russell*, pp. 123–153. Evanston: Northwestern University Press. Reprinted in Feferman *et al*. (1990).

Halbach, V., 2011. *Axiomatic Theories of Truth*. Cambridge: Cambridge University Press.

Heck, R., 1997. “Tarski, truth, and semantics.” *Philosophical Review* 106: 533–554.

Higginbotham, J., 1989. “Knowledge of reference.” In A. George (ed.), *Reflections on Chomsky*, pp. 153–174. Oxford: Basil Blackwell.

———, 1992. “Truth and understanding.” *Philosophical Studies* 65: 3–16.

Horwich, P., 1990. *Truth*. Oxford: Basil Blackwell.

Joachim, H. H., 1906. *The Nature of Truth*. Oxford: Clarendon Press.

Larson, R. and G. Segal, 1995. *Knowledge of Meaning*. Cambridge: MIT Press.

Leeds, S., 1978. “Theories of reference and truth.” *Erkenntnis* 13: 111–129.

Lepore, E. and K. Ludwig, 2005. *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford: Oxford University Press.

———, 2007. *Donald Davidson's Truth-Theoretic Semantics*. Oxford: Oxford University Press.

McGee, V., 1991. *Truth, Vagueness, and Paradox*. Indianapolis: Hackett.

Moore, G. E., 1953. *Some Main Problems of Philosophy*. London: George Allen and Unwin.

Neale, S., 2001. *Facing Facts*. Oxford: Oxford University Press.

Patterson, D., 2012. *Alfred Tarski: Philosophy of Language and Logic*. New York: Palgrave Macmillan.

Putnam, H., 1981. *Reason, Truth and History*. Cambridge: Cambridge University Press.

———, 1985–86. "A comparison of something with something else." *New Literary History* 17: 61–79. Reprinted in Putnam (1994).

———, 1994. *Words and Life*. Cambridge: Harvard University Press. Edited by J. Conant.

Quine, W. V. O., 1953. "Three grades of modal involvement." In *Proceedings of the XIth International Congress of Philosophy*, volume 14, pp. 65–81. Brussels. Reprinted in Quine (1966).

———, 1960. *Word and Object*. Cambridge: MIT Press.

———, 1966. *The Ways of Paradox and Other Essays*. New York: Random House.

———, 1970. *Philosophy of Logic*. Cambridge: Harvard University Press.

Ramsey, F. P., 1927. "Facts and propositions." *Aristotelian Society Supp. Vol.* 7: 153–170. Reprinted in Ramsey (1931).

———, 1931. *The Foundations of Mathematics and Other Essays*. London: Routledge and Kegan Paul.

Rorty, R., 1986. "Pragmatism, Davidson, and truth." In E. Lepore (ed.), *Truth and Interpretation*, pp. 333–355. Oxford: Basil Blackwell.

Russell, B., 1912. *The Problems of Philosophy*. London: Oxford University Press.

Soames, S., 1984. "What is a theory of truth?" *Journal of Philosophy* 81: 411–429.

———, 1992. "Truth, meaning, and understanding." *Philosophical Studies* 65: 17–35.

———, 1999. *Understanding Truth*. Oxford: Oxford University Press.

Strawson, P. F., 1950. "Truth." *Aristotelian Society Supp. Vol.* 24: 129–156. Reprinted in Strawson (1971).

———, 1971. *Logico-Linguistic Papers*. London: Methuen.

Tarski, A., 1935. "Der Wahrheitsbegriff in den formalizierten Sprachen." *Studia Philosophica* 1: 261–405. References are to the translation by J. H. Woodger as "The concept of truth in formalized languages" in Tarski (1983). Original Polish version published in 1933.

———, 1944. “The semantic conception of truth.” *Philosophy and Phenomenological Research* 4: 341–375.

———, 1983. *Logic, Semantics, Metamathematics*. 2nd edition. Indianapolis: Hackett. Edited by J. Corcoran with translations by J. H. Woodger.

Vision, G., 2004. *Veritas: The Correspondence Theory and Its Critics*. Cambridge: MIT Press.

Walker, R. C. S., 1989. *The Coherence Theory of Truth*. London: Routledge.

Williams, M., 1986. “Do we (epistemologists) need a theory of truth?” *Philosophical Topics* 54: 223–242.

Wright, C., 1976. “Truth-conditions and criteria.” *Aristotelian Society Supp. Vol.* 50: 217–245. Reprinted in Wright (1993).

———, 1992. *Truth and Objectivity*. Cambridge: Harvard University Press.

———, 1993. *Realism, Meaning and Truth*. 2nd edition. Oxford: Blackwell.