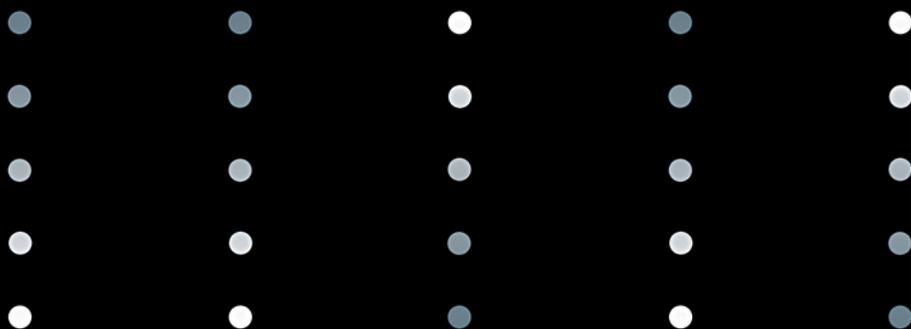


..... ROUTLEDGE STUDIES IN TWENTIETH CENTURY PHILOSOPHY .....

# Donald Davidson

Truth, meaning and knowledge



Edited by Urszula M. Żegleń



**Also available as a printed book  
see title verso for ISBN details**

# Contents

<i>Acknowledgements</i>	vii
<b>Introduction to reading Davidson</b> KIRK LUDWIG AND URSZULA M. ŻEGLĘŃ	1
<b>PART I Polemics</b>	14
<b>1 Is truth a goal of inquiry? Discussion with Rorty</b> DONALD DAVIDSON	15
<b>2 Davidson on deflationism</b> PAUL HORWICH	18
<b>PART II Meaning, truth and interpretation</b>	23
<b>3 Theories of meaning, truth and interpretation</b> KIRK LUDWIG	24
<b>REPLY TO KIRK LUDWIG</b> DONALD DAVIDSON	43
<b>4 How a truth theory can do duty as a theory of meaning</b> GABRIEL SEGAL	45
<b>REPLY TO GABRIEL SEGAL</b> DONALD DAVIDSON	53
<b>5 Radical interpretation and compositional structure</b> PETER PAGIN	55
<b>REPLY TO PETER PAGIN</b> DONALD DAVIDSON	68
<b>PART III Semantics and its applications</b>	71
<b>6 From semantics to ontology, via truth, reference and     quantification</b> STEPHEN NEALE	72
<b>REPLY TO STEPHEN NEALE</b>	82

	DONALD DAVIDSON	
<b>7</b>	<b>Semantics for quotation</b>	<b>85</b>
	HERMAN CAPPELEN AND ERNIE LEPORE	
	<b>REPLY TO HERMAN CAPPELEN AND ERNIE LEPORE</b>	<b>96</b>
	DONALD DAVIDSON	
<b>8</b>	<b>Samesaying</b>	<b>97</b>
	REINALDO ELUGARDO	
	<b>REPLY TO REINALDO ELUGARDO</b>	<b>114</b>
	DONALD DAVIDSON	
<b>PART IV</b>	<b>Knowledge and mind</b>	<b>116</b>
<b>9</b>	<b>McDowell on Quine, Davidson and epistemology</b>	<b>117</b>
	ROGER F.GIBSON	
	<b>REPLY TO ROGER F.GIBSON</b>	<b>128</b>
	DONALD DAVIDSON	
<b>10</b>	<b>Davidson and the new sceptical problem</b>	<b>130</b>
	ANITA AVRAMIDES	
	<b>REPLY TO ANITA AVRAMIDES</b>	<b>148</b>
	DONALD DAVIDSON	
<b>11</b>	<b>General comments</b>	<b>151</b>
	DONALD DAVIDSON	
	<i>Bibliography</i>	<b>155</b>
	<i>Index</i>	<b>174</b>

# Acknowledgements

Donald Davidson's philosophy has achieved its success not only by its new and inspiring ideas, but also thanks to its openness to criticism and discussion, which adds fresh insight to his thought, allows for revision of his views and inspires further inquiries. The current collection is just such a vivid example of doing philosophy. The preparation of this work was possible thanks first of all to Professor Donald Davidson, who acquainted himself with the essays and entered into a discussion of his views. Professor Davidson has also provided a complete bibliography of his writings, which will be very useful for anyone studying his philosophy. I think that I can express my special thanks to Professor Davidson on behalf of all potential readers of this book. I am also very grateful to Kirk Ludwig, who has been assisting me in my work on Davidson's philosophy for a long time. He deserves special thanks for his contribution to this volume. My grateful thanks are due to my friend Tadeusz Szubka who helped me very much and in many ways in preparing this book. I owe thanks to the editors at Routledge, Adrian Driscoll and Tony Bruce, and to the anonymous referees of the manuscript. I also wish to thank my colleagues from Nicholas Copernicus University in Toruń (Poland), Marek Nasieniewski and Scott Thompson, and of course all the contributors, whose deliberations have determined the content of the book. Finally I should express my thanks to Ernie Lepore, who first encouraged me to undertake this project.

Urszula M. Żegleń

## 6

# From semantics to ontology, via truth, reference and quantification

*Stephen Neale*<sup>1</sup>

If we have the semantics of a language right, the objects we assign to the expressions of the language must exist.

(Donald Davidson)

### **Preliminary remarks on reference and ontology**

It would be natural to say that the names ‘Maurice’ and ‘London’, as they occur in an utterance (or inscription) of the sentence ‘Maurice’s mother lived near London’ *refer* to things (a person and a place, respectively), and that someone using these words would be so referring in normal circumstances. It would also be natural to say that the occurrence of ‘Maurice’s mother’ in an utterance (or inscription) of the sentence refers to something (although wrong *philosophically* if Russell is right), and that someone using these words would be so referring. In sharp contrast, it would be decidedly unnatural to say that any of ‘lived’, ‘near’, ‘near London’, ‘lived near London’, or ‘Maurice’s mother lived near London’ referred to things, or that someone using these words would be so referring. But of course this will not prevent philosophers and linguists from arguing that verbs, adjectives, prepositions, prepositional phrases, verb phrases and even sentences do refer to things. A *theoretical* notion of reference may have to replace our ordinary notion as far as a theory of meaning is concerned. And it is no news that the recent history of semantics is littered with theories according to which functions, sets, properties, types, propositions, states of affairs, situations and facts serve as the references of all manner of expression.

From the point of view of constructing a systematic and comprehensive semantics for a natural language, it is not an intuitive or obvious requirement that a final theory appeal to the idea that anything other than names and other singular terms have *references*. It makes no sense to stipulate in advance of detailed investigation that any adequate theory will, or will not, treat, say, sentences as referential devices. Our ordinary uses of words such as ‘refer’, ‘reference’, and so on suggest that it would be at least a little unnatural to view sentences (or even sentences relative to contexts) as referring to things (except in

the sense in which we say, loosely, when examining a text that a particular sentence or verse refers to, say, a recent historical event). But of course the semanticist who appeals to the idea that sentences have references does not have to be entirely shackled by ordinary usage. Certainly we do well to begin with ordinary usage and depart from it grudgingly; but it is well attested that although (many of) our theoretical terms (in both the sciences and the humanities) derive from ordinary usage, as theorising becomes more complex, such terms may begin to diverge in meaning from their ordinary language counterparts, a normal pattern of linguistic development in any interesting discipline.

Donald Davidson (1984a; 1990e) has argued (a) that we can go about constructing theories of meaning that do not assign references to anything other than our standard referential devices (singular terms, including variables under assignment); (b) that such theories will take *objects* and *events*, construed as particulars, to be the sorts of things to which singular terms refer; (c) that as far as the theory of meaning is concerned, we can free ourselves of the idea that predicates refer to *properties* and the idea that individual sentences refer to, stand for, correspond to or are made true by particular *facts* (or by particular *states of affairs* or *situations* that obtain). In addition, Davidson has presented a ‘slingshot’ argument designed to show (d) that even if we wanted facts to serve as the references of true sentences, we could not have them; and he has argued (e) that if there are no facts, then we cannot make sense of ‘correspondence’ theories of truth, for such theories are built upon the idea that a sentence is true if and only if there is some fact to which it corresponds.

### Meaning theory, truth theory and truth definition

A very useful way of characterising Davidson’s approach to meaning is as follows (Davidson does not state it quite like this, but I think what I am about to say is equivalent in all crucial respects to what he says). Point (a) is that a theory of *truth* for a language *L* is any theory that entails, for each sentence *S* of *L*, a true theorem of the form

$$(1) S^* \text{ is true (in } L) \leftrightarrow p$$

a sentence in a metalanguage *M* used to talk about *L* in which ‘*S\**’ is a structural descriptive name (in *M*) of *S* and ‘ $\leftrightarrow$ ’ is the material biconditional. (For the moment, let us ignore the existence of indexical and other context-sensitive expressions.) Instances of (1) are called ‘T-Sentences’. Point (b) is that a theory of *meaning* for a language *L* is, for Davidson, *any* theory, knowledge of which would suffice for interpreting the utterances of *L*-speakers (made in *L*, of course). So (c) a theory of truth  $\theta$  for a language *L* will qualify as a theory of meaning for *L* if knowledge of what  $\theta$  states would suffice for understanding the utterances of *L*-speakers. We can call this the *interpretation* requirement on truth theories. Not every theory of truth for *L* will satisfy it. Imagine a theory of truth  $\theta$

for French that delivered just the following T-sentence for the sentence ‘La neige est blanche’:

(2) ‘La neige est blanche’ is true (in French) ↔ Milan is north of Rome

Since the left- and right-hand sides of this biconditional are both true, (2) itself is also true. Knowing (1) would not suffice for understanding utterances of ‘La neige est blanche’. Taken by itself, this fact is of no import. But in the context of  $\theta$  it has some bite. For  $\theta$ , although delivering a true theorem of the form given in (1) for every sentence of French (including the sentence ‘la neige est blanche’), would not qualify as a theory of meaning for French because knowing  $\theta$  would not provide the means of interpreting utterances of ‘La neige est blanche’. It is this fact that renders  $\theta$  inadequate as a theory of meaning although it is impeccable as a theory of truth.

One way of constructing a theory of truth that might seem to qualify as a theory of meaning would be to stipulate that in every T-sentence, the sentence of  $M$  that replaces ‘ $p$ ’ must have the same meaning as the sentence that replaces ‘ $S$ ’. A theory of truth satisfying this condition would be what Tarski (1956a) called a (materially adequate) *truth definition* for  $L$  (a definition of ‘true-in- $L$ ’). Tarski could help himself to the notion of sameness of meaning (or translation): his objective was to define (i.e. characterise) *truth* (in  $L$ ), not meaning. But an appeal to sameness of meaning (or to translation) is not available to Davidson, nor does he claim it is: Davidson’s objectives are (a) to say *what a theory of meaning is*, and (b) to show how one might go about constructing one, and to do all of this without appealing to ‘meanings’, whether of whole sentences or of their parts. The crucial question, then, is whether it is possible to construct a theory of meaning for  $L$  without appealing to, say, sameness of meaning.

Davidson claims it can be done: a truth theory  $\theta$  for  $L$  will qualify as a theory of meaning for  $L$  if certain empirical and formal constraints are imposed on the way  $\theta$  is constructed. The precise details of the constraints are irrelevant for present concerns. What *is* relevant is that the joint force of the constraints is supposed to ensure that knowledge of  $\theta$  (together with the knowledge that  $\theta$  is a truth theory for  $L$ ) would suffice for understanding utterances of sentences of  $L$ . To speak roughly, the sentence replacing ‘ $p$ ’ in (T) will always (it is hoped) have ‘the same meaning’ (intuitively speaking) as the sentence replacing  $S$ ; more carefully, *there is nothing more to being able to interpret the utterances of  $L$ -speakers than knowing a truth theory constructed in accordance with the constraints in question.*

There is no circularity in this proposal: it is no part of either the truth theory itself or the constraints imposed upon the construction procedure that the sentence replacing ‘ $p$ ’ gives the meaning of, or constitutes, a translation of  $S$ . (Thus an interesting difference between the projects of Davidson and Tarski: whereas Tarski can take the undefined semantic notion of translation for granted and use it to place constraints on what counts as an adequate definition of truth

for  $L$ , Davidson uses a primitive notion of truth in order to characterise what he takes to be the most plausible way of constructing a theory of meaning.)

### Reference, quantification and ontology

I want now to explain a *double asymmetry* between sentences and singular terms in Davidson's programme.

Notice that no appeal to the notion of *reference* (or *predication*, for that matter) has surfaced in the discussion of what a Davidsonian theory of meaning will *look like*. A theory's deliverances with respect to *whole sentences* are all that matter: it is irrelevant how the internal workings of the theory treat the parts of sentences, as long as things come out right for the wholes:

Since T-sentences say nothing whatever about reference, satisfaction, or expressions that are not sentences, the test of the correctness of the theory is independent of intuitions concerning these concepts. Once we have the theory, though, we can explain the truth of sentences on the basis of their structure and the semantic properties of the parts.

(Davidson 1990e, p. 300)

In principle, then, it might be possible to construct an adequate theory that doesn't utilise a notion of, say, reference at all. Only attempts to build adequate truth theories will tell.

According to Davidson, a theory of meaning for a language such as English will, in fact, utilise a *lean*, formal notion of reference, by virtue of containing axioms capable of handling quantification. If we were dealing with an infinite extensional language  $L$  containing only (a finite number of) names, predicates and sentence connectives, but no quantifiers, it would not be necessary to invoke a concept like reference: a finite theory consisting of one axiom for each atomic sentence and one recursive axiom for each sentence connective would suffice because  $L$  would contain only finitely many *atomic* sentences. So it is not the existence of singular terms *per se* that foists reference upon us. According to Davidson, it is quantification that forces us to take something like reference seriously, because it is quantification that forces us to abandon the construction of a straightforward theory whose axioms take the form

(3)  $\_ \text{ is true} \leftrightarrow \dots$

in favour of a theory that takes a detour through satisfaction ('a generalised form of reference'), a theory whose predicate and quantifier axioms take the form

(4)  $(\forall s)(s \text{ satisfies } \_ \leftrightarrow \dots s \dots)$

where 's' ranges over (infinite) sequences of objects.

According to Davidson, the need for such axioms is ontologically significant. The ‘logical form’ of a sentence  $S$  belonging to a language  $L$  is, for Davidson, the structure imposed upon  $S$  in the course of providing an adequate truth theory for  $L$  as a whole; and a truth theory is adequate only if it satisfies the interpretation requirement. Detailed empirical work on the syntax and semantics of natural language suggests that we will not get very far in our attempts to construct an adequate truth theory for English unless we view the logical forms of English sentences as encoding something very like the quantifier-variable structures familiar from formal languages such as the first-order predicate calculus. One formally useful way of pulling together names and variables (and other singular terms, if there are any) within a truth theory—a notational variant of many other ways, and perfectly consonant with Davidson’s approach—is to use the notion of reference relative to a sequence, which we can abbreviate as ‘*Ref*’. On this account, the axiom for the name ‘Maurice’ is given not by a simple reference axiom such as (5) but by (6):

(5) The referent of ‘Maurice’=Maurice

(6)  $(\forall s)(Ref\langle \text{‘Maurice’}, s \rangle = \text{Maurice})$

where ‘ $s$ ’ ranges over sequences. Axiom (6) is a ‘reference’ axiom in the sense that *it connects a piece of language with some entity or other, some piece of the world*. (For ease of exposition, let us put aside names that allegedly fail to refer, if there are such expressions.) Axioms for individual variables are given by the axiom schema (7):

(7)  $(\forall s)(\forall k)(Ref\langle \Gamma x_k \bar{,} s \rangle = s_k)$

where ‘ $s$ ’ ranges over sequences and  $s_k$  is the  $k$ -th element of  $s$ . (The difference between a name and a variable, then, is that the *Ref* of a name is always the same entity, whereas the *Ref* of a variable depends upon the sequence in question. The constant/variable distinction should not be confused with the rigid/non-rigid distinction: variables (with respect to sequences) are just as rigid as names (in Saul Kripke’s sense: 1980).

Two interesting things about reference axioms should be noted: (a) they pair bits of language with bits of the world (this gives them their ontological bite); and (b) the fact that reference is no more than such a pairing—any pairing that works—means that there is no need to provide an analysis of the reference relationship. Questions about whether Davidson’s story about reference conflicts with those of, say, Searle 1983; Donnellan 1971; Kripke 1980; or Evans 1982, don’t make a great deal of sense.

Axioms for verbs, adjectival expressions and ordinary common nouns treat such expressions as predicates (for ease of exposition, let us put aside so-called ‘intensional’ predicates such as ‘fears’, ‘fake’, and so on).<sup>2</sup> For example:

- (8)  $(\forall s \forall \alpha)(s \text{ satisfies } [\alpha \text{ snores}] \leftrightarrow \text{Ref}(\alpha, s) \text{ snores})$   
 (9)  $(\forall s \forall \alpha \forall \beta)(s \text{ satisfies } [\alpha \text{ likes } \beta] \leftrightarrow \text{Ref}(\alpha, s) \text{ likes } \text{Ref}(\beta, s))$

The utility of the axioms flowing from (7) lies in their interaction with axioms for quantifiers. If quantification in English turns out to be unrestricted and completely analysable in terms of the first-order quantifiers ‘ $\forall$ ’ and ‘ $\exists$ ’—a more realistic proposal will be considered shortly—a truth theory for English can get by with an axiom like (10), and its universal counterpart:

- (10)  $(\forall s)(\forall k)(\forall \Phi)(s \text{ satisfies } [(\exists x_k)\Phi] \leftrightarrow \text{there is at least one sequence differing from } s \text{ at most in the } k\text{-th place that satisfies } \Phi)$

The theory is completed by adding that a sentence is true if and only if it is satisfied by every sequence.

Assuming an adequate background logic (e.g. extensional, first-order logic with identity), we could then prove theorems (T-sentences) such as the following:

- (11) ‘Maurice is Greek’ is true  $\leftrightarrow$  Maurice is Greek  
 )  
 (12) ‘Maurice likes everything’ is true  $\leftrightarrow$  Maurice likes everything<sup>3</sup>  
 )

If it turns out that a theory of truth that is to succeed as a theory of meaning cannot be constructed without making use of some notion of reference (such as *Ref* above), i.e. without axioms that connect pieces of language with non-linguistic entities, let us say that such a notion is ‘theoretically ineliminable’.

According to Davidson, accepting that reference is theoretically ineliminable does not mean accepting that *any particular set of reference axioms* is ineliminable. Indeed, it is Davidson’s position, as it is Quine’s, that it is possible to transpose any adequate set of truth-theoretic axioms *X* into another adequate set *Y* that contains, as a subset, a set of reference axioms quite different from those contained in *X* (the predicate axioms would also differ, of course).<sup>4</sup> A particular axiomatisation is tested only by its proper theorems, i.e. by its deliverances at the level of whole sentences. So the notion of reference employed by Davidson is philosophically lean in two senses: (a) no particular set of reference axioms is privileged; and (b) to claim that a well-understood notion of reference is appealed to by Davidson in constructing a theory of meaning for a language is not to claim very much; for on Davidson’s account to say that ‘Maurice’ refers to Maurice is just to say that there is a successful axiomatisation containing the axiom (6) above—or some notational variant that hooks up ‘Maurice’ and Maurice—where ‘*Ref*’ is straightforwardly an extensional functor or a description-forming operator.

Two points need to be made in this regard. First, as Davidson (1993b.) stresses, the fact that an adequate truth theory might make an ‘unnatural’

assignment of objects to individual words does not affect the overall ontology to which the language is committed. (Of course, no *successful* assignment is really ‘unnatural’ on Davidson’s account.) Second, no appeal to modes of presentations, causal chains, informational packages, or intentionality is needed in order to characterise the theoretical notion of reference that Davidson employs. As Richard Rorty observes, for Davidson ‘any ‘theory of truth’ which analyses a relation between bits of language and bits of non-language is already on the wrong track’ (Rorty 1986, p. 333).

There is a crucial difference between the reference axioms and the predicate axioms in a successful truth theory. The former assign particular entities (individuals) as the semantical values (references) of expressions; the latter do no such thing. This is important. On Davidson’s account, since reference axioms are theoretically ineliminable, we must accept the entities that our singular terms stand for, even though we do not have to regard any particular reference function as privileged. Predicate axioms invoke no new entities: the only ontologically significant notion they use is satisfaction, and satisfiers are just sequences of objects. By hypothesis, a successful axiomatisation construes singular terms as *standing for* objects, and construes sentences as satisfied by sequences of objects. Hence Davidson’s ‘realism’ about ‘the familiar objects whose antics make our sentences and opinions true or false’ (Davidson 1984a, p. 198).

On Davidson’s account, *events*—taken to be unrepeatable particulars—will get into the picture along with *objects* because a successful axiomatisation will have to deal with sentences which involve quantification over events, for example (13) and (14):

- (13 There was a fire and there was a short-circuit  
 )  $(\exists x)\text{fire } x \ \& \ (\exists y)\text{short-circuit } y$   
 (14 There was a fire because there was a short-circuit  
 )  $(\exists x)(\text{fire } x \ \& \ (\exists y)(\text{short-circuit } y \ \& \ y \text{ caused } x))$

Sentences containing action verbs and adverbs (‘John left quickly’) and those containing bare infinitives (‘John saw Mary leave’) also appear to require quantification over events.<sup>5</sup> In addition to claiming that we will need to pair bits of language with *objects* and *events* as far as constructing a theory of meaning is concerned, Davidson suggests that if the need to posit a particular ontological category does not arise in the construction of an interpretive truth theory, then the need cannot arise at all. (Of course, linguistic categories and set theoretic entities such as sequences are posited by the metalanguage.) The thought behind this suggestion appears to be (roughly) the following (although I have not actually found it stated quite this way anywhere in Davidson’s work): A theory of truth for *L*, of the sort that can serve as a theory of meaning for *L*, delivers a true theorem of the requisite form for *every* sentence of *L*. So there is nothing one can say in *L* that outstrips the ontology revealed by an interpretive truth theory for *L*; so there is no sense to be made of ontological categories not forced upon us by

the construction of a viable semantics. (It might be objected that in doing semantics we can appeal only to entities that we think exist, so a semantic theory offers us no more by way of ontological insight than ordinary reflection upon our thought and talk. But this misses the fact that an adequate semantics is *systematic* in ways in which ordinary reflection is not.)

One interesting question that is left open concerns competing truth theories with different ontologies. If  $\theta_1$  posits *As*, *Bs* and *Cs*, while  $\theta_2$  posits only *As* and *Bs*, then if both are adequate truth theories for *L*, we have reason to posit only *As* and *Bs*. In this case our ontology is given by the intersection of the things posited by the competing theories. But what if  $\theta_2$  posited only *Bs* and *Cs*? Perhaps it is unlikely that we will find ourselves in such a situation (we are unlikely to find even *one* theory that covers all the data), but the question is still of philosophical interest, if we take seriously the idea that ontology flows from semantics in the way that Davidson suggests. Presumably Davidson will prefer a theory that posits objects and events over one that posits (for example) events and properties on the grounds that (a) identity conditions for objects and events are considerably clearer than they are for properties, and (b) our best accounts of nature and our most cogent statements of traditional philosophical problems concerning (for example) causation, time, change, human action, and the mind-body problem appear to presuppose the existence of objects and events.

### Potential challenges and modifications

I have tried only to present clearly the bare bones of Davidson's ambitious programme. I want to conclude with a partial catalogue of ontological issues that could, in principle, have a bearing on the final shape of an interpretive truth theory.

Davidson proposes an elegant semantics and a tidy ontology of objects and events. There are no facts, no situations, no states of affairs, no propositions and no properties because we can construct interpretive truth theories without positing them. Of course Davidson does not claim to have a *proof* that we need only objects and events. His claim is simply that there is no good evidence to date that we need posit more than objects and events. In a way, Davidson is throwing down the gauntlet: 'Show me sentences that appear to require more than objects and events and I will do my best to show that objects and events suffice.'

*Context-sensitivity* Truth-theoretic Axioms for words such as 'I', 'me', 'we', 'us', 'you', 'he', 'him', 'she', 'her', 'it', 'they', 'them', 'this', 'these', 'that', 'those', 'today', 'yesterday', 'tomorrow', 'here', 'there', 'now', 'then', 'hitherto', 'henceforth', 'present', 'current', 'contemporary', 'previous' and 'next' must take into account contextual features of one sort or another. In view of this, the Davidsonian will naturally treat truth as a property of utterances rather than sentences, and then treat utterances as individual events. Simplifying somewhat, on such an account, axioms will be rewritten with metalanguage

semantic predicates taking on a parameter for the utterance event  $e$ , giving us, for example, (15) for ‘Maurice’ and (16) for ‘I’:

- (15  $(\forall s)(\forall e)(Ref('Maurice', s, e)=Maurice)$   
 )  
 (16  $(\forall s)(\forall e)(Ref('I', s, e)=e_u)$   
 )

where  $e_u$  is the utterer, the person producing the utterance (e.g. the speaker or writer).<sup>6</sup>

*Properties* A sentence like (17) might be thought to present Davidson with an interesting challenge as it appears to involve quantification over colours, construed as properties:

- (17 This is the same colour as that  
 )

The Davidsonian who does not want an ontology of properties must either find a way of understanding colours as things other than properties, or else show that once the logical form of (17) is revealed, there is no quantification over colours.

*Facts* If we were to find that we needed to allow variables to range over, say, *facts* or *situations* in order to provide an adequate truth theory, then on Davidson’s account, facts or situations would also be part of our ontology. But Davidson suggests (a) that the need will not arise, and (b) that entities like facts and situations (under their most common construals) are ruled out independently, as can be demonstrated by a ‘slingshot’ argument.

In the case of (a) the sorts of sentence that might tempt one to posit facts include the following:

- (18 The fact that there was a short-circuit caused it to be the case that there was  
 ) a fire

Davidson suggests that (12) above gives the logical form of (18) so an appeal to facts is unnecessary. A more difficult case for Davidson might be:

- (19 The fact that Mary left Bill’s party did not worry him, but the fact that she left  
 ) so suddenly did

Davidson is surely correct that events, just like objects, can be described in a myriad of ways. If Mary’s leaving was Mary’s sudden leaving, then how is it that (19) can be true? A promising approach to this problem utilises Grice’s (1989) notion of conventional implicature in an attempt to provide the basis of an explanation of our judgements concerning (19).

With regard to (b) Davidson (1984a; 1990e; 1993b) argues that we cannot have facts of any significance because of a simple formal argument that has come to be called the ‘slingshot’ and which can be cast in a formal mode by (i) assuming that ‘the fact that...caused it to be the case that...’ is a non-truth-functional sentence connective and then (ii) deriving the conclusion that it is truth-functional. I have discussed this type of argument in detail elsewhere (Neale 1995; Neale and Dever 1997). As far as facts are concerned, my main conclusion may be summarised thus: slingshot arguments do not actually *rule out* facts but they do impose very definite constraints on what theories of facts must look like, constraints that many proposed theories are incapable of satisfying. (Russell’s theory of facts passes—assuming his Theory of Descriptions; the theories of Wittgenstein and Austin fail.)

*Modality and possible worlds* A truth theory for English must be able to handle sentences that contain modal adverbs such as ‘necessarily’ and ‘possibly’. Will this require an ontology of ‘possible worlds’? If the axioms of a truth theory are modalised, a richer background logic (a suitable modal logic) can be used to derive theorems of the requisite form while treating ‘necessarily’ and ‘possibly’ as non-extensional sentence connectives (Peacocke 1978). Of course, Davidson himself has little time for such connectives, but in view of the fact that Quine’s worries about the modalities were overstated, and the fact that slingshot proofs do not actually finish off modal sentence connectives, there is good reason to think that possible worlds are not required.

*Higher-order quantification* It is well known that natural language quantification is more complex than it is in standard first-order languages. However, work in generative linguistics and mathematical logic has revealed elegant methods of extending Tarski’s insights, so that truth theories can be provided for quantified fragments of natural languages while making precise the relationship between the superficial grammatical form of a sentence and its logical form in Davidson’s sense (Wiggins 1980). Many of the details need not concern us here. Suffice to say that noun phrases such as ‘some man’, ‘every Athenian’, ‘no farmers’, ‘most tall soldiers’, etc. can be viewed as restricted quantifiers composed of quantificational determiners (‘every’, ‘some’, ‘no’, ‘most’, etc.) combined with simple or complex nouns (‘man’, ‘tall man’, etc.). Axioms such as the following make things run very smoothly:

(20)  $(\forall s \forall k \forall \Phi \forall \psi)(s \text{ satisfies } \llbracket \text{some } x_k: \Phi \rrbracket \psi \leftrightarrow \text{some sequence satisfying } \Phi$   
 ) and differing from  $s$  at most in the  $k$ -th place also satisfies  $\psi$ )

(21)  $(\forall s \forall k \forall \Phi \forall \psi)(s \text{ satisfies } \llbracket \text{most } x_k: \Phi \rrbracket \psi \leftrightarrow \text{most sequences satisfying } \Phi$   
 ) and differing from  $s$  at most in the  $k$ -th place also satisfies  $\psi$ )

It is not necessary to pack any additional set theory or overtly higher-order machinery into the right-hand sides of (20) and (21); they have the same form as axioms for the traditional unrestricted quantifiers *modulo* the relevant restriction concerning the satisfaction of  $\Phi$ .

### Notes

- 1 I thank Donald Davidson, Ernie Lepore, Peter Pagin and Bruce Vermazen for comments on ancestors of this paper.
- 2 There are notorious difficulties involved in providing a uniform predicational analysis of these categories. For present purposes, the differences between these categories can be ignored as they do not raise problems that bear directly on my main theme.
- 3 To be precise, we could prove such theorems using extensional first-order logic with identity *if* we treated 'Ref ("Clinton",s)' either (a) as a Russellian definite description ('the referent of "Clinton" with respect to s') and hence as a first-order definable device of quantification; or (b) as a complex singular term formed from a singular term and a functional expression. If method (a) is selected, proofs will make use of Whitehead and Russell 1927 \*14.15, a derived rule of inference for extensional contexts, rather than straightforward applications of the Principle of Substitutivity. If method (b) is selected, a version of first-order logic with functors must be selected. There are, I believe, reasons for preferring method (a).
- 4 It is debatable whether interpretive truth theories for languages containing Russellian descriptions together with words such as 'necessarily' and 'possibly' allow of the sorts of straightforward reference permutations Davidson has in mind. This does not matter for present concerns.
- 5 On these matters, see Davidson 1980a and Higginbotham 1983.
- 6 For sophisticated and more plausible implementations, see (e.g.) Burge 1974; Weinstein 1975; Taylor 1980.

### REPLY TO STEPHEN NEALE

*Donald Davidson*

Neale gives a masterful and sympathetic account of my views on the sort of theory I favour as the basis for interpreting a speaker, and in particular of the relation between such a theory and ontology. He also raises a number of problems or challenges that arise in the attempt to apply such a theory to a natural language with normal expressive power.

One such problem he mentions only suggests that perhaps the constraints on a satisfactory theory can solve it. Since Paul Horwich also mentions this problem (which he apparently thinks insoluble), let me say a few words about it. The problem is how to eliminate such T-sentences as

(1) 'Snow is white' is true ↔ grass is green

or Neale's (2). I have from time to time made suggestions on how to cope with this difficulty, not all of them consistent with each other. My present view is that there is no point in trying to find a strictly formal solution. What is wrong with (1) is that it will not fit into an empirically satisfactory theory because it will lose track of relations among sentences. Simple inferences such as that from 'This is

snow' and (1) to 'This is white' will not be validated by our theory, since it must treat 'snow' as ambiguous. (There is nothing wrong with (1) or Neale's (2) if they are embedded in theories which systematically alter the extensions of all predicates and names in such a way as to preserve all inferences. Those who raise this difficulty don't have such cases in mind; they think of the rest of the theory as providing the 'standard' translations.) The thing to bear in mind is that a theory of truth, as I want to employ it, is an empirical theory, and is constructed in the light of certain intuitively obvious kinds of evidence. These sources of evidence are mainly three: a matching of perceptual sentences like 'This is snow', 'That is white' to their socially identified causes; the logical relations inferred to hold among sentences; and relations of evidential support among sentences. (I have spelled out how I think this can work in 'The Structure and Content of Truth': 1990e.) I now address the challenges Neale sees to my relatively simple semantics.

*Context-sensitivity* I recognised from the start that the truth conditions of many sentences are sensitive in systematic ways to context, and so a theory of truth for a natural language would have to treat utterances as the elements with which the theory treats. I made a simple suggestion of how this might go in 'Truth and Meaning', and my then-student Scott Weinstein made a far more sophisticated proposal. No doubt there is more work to be done.

*Properties* Neale seems to me to confuse two issues about properties. The first is whether there is in general any objection to including properties in our ontology. The second is whether there is any advantage in introducing predicates as the sole semantic item used to explain the function of predicates. I have no objection to accepting properties, if they turn out to be needed to explain such sentences as 'This is the same colour as that'. I have no general objection to abstracta, as long as they are shown to be useful; as I always say, they take up no space. I have deep misgivings, however, about semantic theories such as Russell's, which treat predicates as referring to, or corresponding to, predicates. These misgivings have two related sources. First, such theories are apt, as Plato discovered, to lead to regresses. The second is that it is essential to my idea of how a theory of truth is verified that the right side of T-sentences not employ conceptual resources not employed by the sentence for which the truth conditions are being given. This constraint is obviously violated in the treatment of indexicals, but not in the basic way it would be violated if the work of every predicate were explained by appeal to a property (relation, etc.).

*Facts* Neale's splendid discussion of the slingshot argument against facts, while it shows the limitations of that argument in both Church's and Gödel's versions (both of which I repeated in various places), has encouraged me to stick to the conclusion that facts cannot be incorporated into a satisfactory theory of truth. Neale speaks of Russell's view of facts as 'acceptable', but he does not show that it can be incorporated into a Tarski-style truth theory.

*Modality and possible worlds* A satisfactory theory of a natural language must say something about modal adverbs, and I have not published anything on the

subject. Neale is right that I am unhappy with non-extensional sentential connectives; I am even more unhappy with analyses that depend on our intuitions about possible worlds. I do think there are ways to explain in a non-intensional way how modal adverbs function, but here is not the place to explore this.

*Higher-order quantification* I can only applaud the suggestion that the semantic functions of expressions such as ‘some man’ and ‘most people’ can be accommodated within a harmlessly expanded version of a theory of truth of the sort I favour.