23

Donald Davidson (1917–)

ERNEST LEPORE

Donald Davidson is one of the most important and influential philosophers of the second half of the twentieth century. He has never attempted a systematic exposition of his philosophical program, so there is no single place a student, interpreter, or critic can seek its official formulation. His published essays, taken together, form a mosaic that must be viewed all at once in order to discern an overall pattern. In addition, they sometimes exhibit an enigmatic quality, with subtleties, complexities, and cross-references that often cannot be entirely appreciated except in conjunction with each other. All of this can render access to his thought not just difficult but at times frustrating, despite its obvious importance.

In an effort to help a novice or even someone already stumped, the following major themes in Davidson's philosophy will be summarized: (1) Reasons and Causes, (2) Events and Causation, (3) Anomalous Monism, (4) Theory of Meaning and Compositionality, (5) Radical Interpretation, (6) Adverbial Modification, (7) The Method of Truth in Metaphysics, (8) Against Facts, (9) Truth and Correspondence, (10) Animal Thought, (11) Alternative Conceptual Schemes, (12) Anti-skepticism, (13) Anti-Cartesianism and First Person Authority, (14) The Rejection of Empiricism.

Though this list is not exhaustive, it identifies broad themes that structure Davidson's project. Of course, nothing short of reading Davidson's own work can supplant a detailed characterization of his take on each individual point, but I hope to tempt you to journey into Davidson's philosophical world.

Reasons and causes

We often try to explain another's actions by citing her reasons for so acting. Mary offers Bill the seat next to her on a train. She does so because she wants Bill to sit down and believes that by identifying this open seat he will. She thereupon asks him to sit down. What relation must obtain between Mary's behavior and her reason for this behavior in order to correctly conclude that she acted as she did *because* of her reason?

Until Davidson's "Actions, Reasons, and Causes" (see 1980), as difficult as it now seems to comprehend, something close to a consensus had formed in philosophy that whatever the relationship might be it could not be causal. It was believed that an alleged "logical connection" between reasons and actions excluded any causal relation

between them. The central purpose of Davidson's essay was "to defend the ancient – and commonsense – position that rationalization is a species of causal explanation" (1980: 3). Much of his essay is devoted to refuting various arguments, then popular, that purported to show that reasons could not cause the actions they rationalize. According to each, a necessary condition for causal interaction cannot be satisfied by reasons and actions. These arguments are too many to be properly treated here, but in passing it should be noted that they were inspired by remarks of the twentieth-century philosopher Ludwig Wittgenstein or by certain interpretations of the eighteenth-century British philosopher David Hume's strictures on causation.

By way of a single example, many authors believe that Hume had more or less established that in order for one thing *A* to be causally related to another thing *B*, *A* and *B* must not be logically connected. So, for example, when one billiard ball moving with a certain momentum hits another, then, unless circumstances are unusual, the second will be caused to move. The second ball's movement occurs not as a matter of logic, but rather as a matter of the physical nature of our universe. It is logically possible that causal interaction in this universe has been governed by a physical nature distinct from what actually governs it. One intuition behind denying that reasons can be causes of actions is that they are logically related; so, if I believe that smoking is harmful, and I desire no harm, mustn't I, as a matter of logic or of meaning alone, intend not to smoke? If so, then how could my reasons for intending not to smoke have caused me to have that intention?

For our purposes, it suffices to say that Davidson replied to this argument and others by showing either that reasons and actions indeed satisfy the necessary condition in question, or that the would-be necessary condition for causal interaction is in fact not one at all. So, with respect to Hume's observation, Davidson argued that a logical connection between descriptions of a cause and an effect does not by itself preempt causation, as is evident in "The cause of event *e* caused event *e*." No one would infer from the fact that an event could not have existed under the description "the cause of *e*" without being the cause of the event *e* that the first didn't cause the second. Similarly, even if no one could be described as having a belief under the description of "smoking is harmful" and having a desire under the description "don't do anything harmful" without having an intention under the description "don't smoke," it doesn't follow that there is no causal relation between these reasons and the intentional action that ensues.

Events and causation

In "The Individuation of Events" (in 1980), Davidson argues that we must recognize that we can describe the same action in different ways. Indeed, if we did not do this, we could not make sense of perfectly natural claims like "Jones managed to apologize by saying 'I apologize'," where a single action is described both as a managing to apologize and as a saying "I apologize." But what sort of object are these actions that admit of re-description?

Tables, chairs, and people are concrete, dated particulars, that is, unrepeatable entities with location in space and time. These features alone distinguish them from, say, either numbers or God. But what about actions and events, for example, the action

of Bob shooting Bill, or the event of Hurricane Floyd, or the event of the stock market crash of October 1929? Anyone who doubts these three events exist need only confer with one of their victims. But what sort of entity can an action or event be? We speak, for example, of the event of a baseball game between the Yankees and the Rangers as lasting three hours, and as taking place in New York. We might conclude that this event was exciting or too long or even crucial in determining a champion. On the basis of on such assessments what can we conclude about what sort of thing an event is?

Davidson's chief claim about actions and events is that like tables and chairs they are concrete, dated particulars that can be described in various nonlogically equivalent and non-synonymous ways. What distinguishes them from other sorts of concrete, dated particulars is their potential for causal interaction, and so, it's part of the nature of being an event that it can stand in a causal relationship. Davidson goes on to individuate them, so that two events are identical just in case they have the same causes and effects. Since Davidson treats causation as a relation between events, and takes action to be but a species of event, events comprise the very subject matter of action theory, as well as science and ethics. (We will take up below his argument for their existence and for specific claims as to their nature in the section on the method of truth in metaphysics.)

Davidson holds that events related by causation must be subsumable under some law or other, where a law is a generalization confirmable by its positive instances and, if true, supports counterfactual statements, and where an event is subsumed by a law just in case it instantiates that law (1980: 217). So, for example, according to Boyle's Law, the pressure of a fixed mass of gas at a constant temperature is inversely proportional to the volume of that gas. An instance of this law – if a is the pressure of a fixed mass of gas at a constant temperature, then a is inversely proportional to the volume of that gas – is *confirmed* just in case when its antecedent is true, so is its consequent. To say it *supports* its counterfactual instances means that even if a is not the pressure of a fixed mass of gas at a constant temperature, if it *were*, then it *would* be inversely proportional to the volume of that gas (1980: 215).

Davidson's views about the nature of events and their relation to laws brought him to a stunning conclusion about the relationship between minds and bodies, namely, his thesis of anomalous monism, to which we turn immediately.

Anomalous monism

Much can and has been said in favor of each of the following three claims:

1   The mental and the physical are distinct.
2   The mental and the physical causally interact.
3   The physical is causally closed.

The problem, though, is that they seem inconsistent. Consider their application to events. (1) says that no mental event is a physical event; (2) says that some mental events cause physical events, and vice versa; a loud noise reaching Tom's ear may cause him a desire to turn down his radio; and his desire to turn down his radio may cause his arm to move in such a way to result in the volume of his radio being lowered. (3) says that all the causes of physical events are themselves physical events. The dilemma

posed by the plausibility of each of these claims and by their apparent incompatibility is the traditional mind–body problem. Davidson's resolution, as articulated in "Mental Events" (in 1980), "The Material Mind" (in 1980), and "Philosophy as Psychology" (in 1980) consists of theses (4)–(6), which taken together comprise his thesis of *anomalous monism*:

4   There are no exceptionless psychological or psychophysical laws, and in fact all exceptionless laws can be expressed in a purely physical vocabulary (1980: 214–15, 231).
5   Mental events causally interact with physical events (1980: 208).
6   Event *c* causes event *e* only if an exceptionless causal law subsumes *c* and *e* (1980: 208).

The thesis is monistic, since it assumes there is but one kind of stuff in the world, physical stuff, but it is anomalous, since although its monism commits it to physical and mental stuff being the same stuff, it denies that that there is a strict reduction of the one to other. A full-blown exegesis and defense of this view and the claims which comprise it is beyond the intended scope of this essay, but a few words about the extra-ordinary thesis (4) are in order.

Thesis (4) is a version of (1). It is commonly held that whatever property a mentalistic predicate M expresses is *reducible* to one expressed by a physical predicate P (where "M" and "P" are not logically connected) only if an exceptionless law links them (where an exceptionless law is exactly what it says, namely, one that under no conditions admits of exceptions. Boyle's Law, stated above, is obviously not an exceptionless law). According to (4), mental and physical properties are distinct. In sketch, Davidson bases his argument against the possibility of exceptionless laws linking mental and physical predicates on their sufficiently distinct *constitutive* principles (1980: 222, 238).

Measurability of length, mass, temperature, and time are *constitutive* of the physical, inasmuch as these features govern the applicability of physical predicates (1980: 221). So, for example, anything physical must have length. Suppose, though, upon investigation, we discover that among three physical items, though the first is longer than the second and the second longer than the third, the first is *not* longer than the third. What would we conclude? We would assume that either we are mistaken, or that their lengths changed during the course of measurement. What we would not conclude is that the transitivity of length is false. Why? Because, no three things could have a physical predicate true of them unless their lengths respect transitivity; that is, respecting this constraint is constitutive of being physical.

With mental items, their constitutive principles include principles of rationality, for example, constraints about consistency and rational coherence (1980: 236–7). For example, the transitivity of desire – if a person desires *a* over *b* and *b* over *c*, he *ought* to desire *a* over *c*; or the consistency of belief – we assume that if an agent believes that *p*, he ought not also to believe that not *p*. This ties in with earlier discussion of reasons and actions; interpreting the behavior of another requires attributing beliefs and desires. These attributions are intended to provide an agent's rationale for acting, but they fail this task unless a degree of *rational* choice is presumed. By virtue of this presumption, the constitutive principles of the mental include *norms* of rationality. Davidson claims such constitutive principles have "no echo in physical theory" (1980:

231). So, he concludes, in an important sense psychology cannot "be reduced to the physical sciences" (1980: 259), namely, exceptionless laws cannot link the two sorts of sciences, because the normative relationships among mental states cannot be expressed in a physical language.

This compact discussion of *anomalousness* leaves a host of questions unanswered, but it serves to identify points of controversy surrounding Davidson's thesis (4). What about his thesis (6)? In "Causal Relations" (in 1980), Davidson argues that the most plausible interpretation of singular causal statements like "The short circuit caused the fire" treats them as two-place predicate statements with their singular terms, in this case, "the short circuit" and "the fire" designating events. Thesis (6) says that an event c causes an event e only if there are singular descriptions D of c and D′ of e, and an exceptionless causal law L such that L and "D occurred" entail "D caused D′" (1980: 158). But (6) and the second part of (4) entail that physical events have only physical causes, and that all event causation is physically grounded.

Given the parallels between (1)–(3) and (4)–(6), it may seem that the latter, too, are incompatible. Davidson, however, argues that they can all be true if (and only if) individual (token) mental events are identical to individual (token) physical ones (1980: 215, 223–4). Suppose an event e is physical just in case e satisfies a predicate of (basic) physical science, where such predicates are those that occur in exceptionless laws. Since Davidson assumes that only physical predicates (or predicates expressing properties reducible to physical properties) occur in exceptionless laws (1984b: 240), it follows that every event that enters into causal relations satisfies a physical predicate. But, then, it follows that a mental event that enters into a causal relation must satisfy some physical predicate, and so, is itself a physical event. His argument, if sound, establishes no more than that every concrete mental event is identical to some concrete physical one. Since this identity of mental and physical event tokens is compatible with rejecting systematic laws bridging mental and physical event-types, that is, with anomalous monism, the latter thesis only partially endorses (1). The mental and physical remain type-distinct insofar as mental and physical events are *not* linked by exceptionless laws under mental and physical descriptions.

Davidson's account of reasons and events impacts on, and is in turn, affected by, his radical approach to language; to see this, we turn now to examine his theory of meaning.

Theory of meaning and compositionality

Our discussion will be limited primarily to ideas present in "Theories of Meaning and Learnable Languages" and "Truth and Meaning" (both in 1984b), where Davidson identifies an adequacy criterion for theories of meaning for natural languages, and then applies it critically to a number of then prominent analyses of aspects of natural language (see TARSKI, CHURCH, GÖDEL). He also sketches a program in which, surprisingly, a certain austere style of meaning theory meets this adequacy criterion.

What aim(s) should a theory of meaning seek to accomplish? That will depend on which linguistic aspects a theorist wants to explain. So, for example, though natural languages are spoken by finite speakers without magical abilities, they still have an infinity of meaningful (non-synonymous) sentences, each of which, at least potentially,

a speaker could understand (at a given time). For any (indicative) sentence S of English, a new one can be formed by prefacing it with "It is believed that." For any two (indicative) sentences, S and S′, a new one can be formed by disjoining them with the word "or"; and so on for other productive mechanisms of our language. The novel sentences which these productive mechanisms give rise to are intelligible to normal speakers if their components are. This capacity seems to require that speakers have learned (a finite number of) rules that determine from a finite set of *semantic primitives* what counts as meaningful compositions (where an expression is semantically primitive if the "rules which give the meaning for the sentences in which it does not appear do not suffice to determine the meaning of the sentences in which it does appear" (1984b: 9)).

On the basis of such considerations, Davidson requires of a theory of meaning that it specify what every sentence means by exhibiting its meaning as a function of the meaning of its significant parts (based, presumably, on their arrangement in the sentence). Let's call any such theory for a language "a compositional meaning theory" for that language. Davidson was the first philosopher to bring to prominence the importance of the requirement that a theory of meaning of our language exhibit it as compositional (1984b: 23). The requirement focuses attention on the need to uncover structure in natural languages. While this is clearly something that philosophers from time immemorial have engaged in, this project had not been, until Davidson, clearly separated from the ancient but now discredited project of conceptual analysis.

Davidson's positive suggestion for a compositional meaning theory for a language L, surprisingly, utilizes no concept of *meaning* that goes beyond truth. To wit, his theory of meaning takes the form of a (finite) theory *of truth* that, for each sentence S of L, entails what we shall call a T-sentence of form

(T)   S is true in L if, and only if, *p*,

where "*p*" specifies (in a metalanguage) conditions under which S is true in L. So, for example, an adequate compositional meaning theory for German should issue in a theorem like (S):

(S)   "Schnee ist weiss" is true in German if, and only if, snow is white.

Why does Davidson choose this rather austere form of theory over, say, one that explicitly invokes meaning by issuing in theorems of the form

(M)   S in L means that *p*,

where "*p*" specifies (in a metalanguage) what S means in L? His reasons have nothing, as sometimes suggested, to do with replacing the complex notion of meaning with one more tractable or easily understood. Davidson's inquiry is not guided by conceptual or metaphysical qualms about the notion of meaning, but solely by the goal of devising a compositional meaning theory. He argues that this aim can be achieved with a compositional meaning theory that issues in theorems that take the form of (T), but not by one that issues in theorems of form (M). (His reasons are based on the fact that unlike the locution "is true if and only if," the locution "means that" is semantically opaque, thus hindering the development of a compositional meaning theory. What semantic opacity is will be discussed below in the section on animal thought.)

A compositional theory of meaning for a language L that issues in interpretive T-sentences like (S) is such that anyone who knows it is positioned to understand every sentence of L. By specifying the meaning of a sentence S in L *via* sentences of form (T), Davidson is requiring a specification that enables anyone who understands the language in which the specification is given (the *metalanguage*) to understand (the object language sentence) S. The observation that natural languages are compositional is the foundation upon which Davidson builds his program in the theory of meaning.

Since we do not and could not know a priori how to interpret the expressions of a natural language or how to assign them interpretive truth conditions, an adequate compositional meaning theory must be *empirical*. Any such theory must function as a theory for a natural language. In the case of one's own language, of course, though what one knows about it is not a priori, no difficulty arises in identifying which sentences of form (T) are interpretive. The problem for a theorist concerning a language he already understands is simply to figure out what the axioms of the theory ought to be in order to construct proofs of them. A quite different problem confronts a theorist for foreign languages (even to some degree for other speakers of one's own language, a theme emphasized more in Davidson's later work). This is where the important notion of *radical interpretation* enters into his discussion.

Radical interpretation

What justifies a choice of one compositional meaning theory for a language over another? Davidson argues that an adequate compositional meaning theory must be empirically warranted under the practice of *radical interpretation*. What this means is that certain specific empirical considerations must be respected in choosing between distinct but true compositional meaning theories, namely, in opting for a compositional meaning theory for German that issues in (S) over one that issues in (W).

(W)   "Schnee ist weiss" is true in German if, and only if, grass is green.

(W) is, as a matter of fact, true, but, unlike (S), it fails to *interpret* "Schnee ist weiss," and so no compositional meaning theory for German that issues in (W) can be adequate. But for languages we do not already understand, a compositional meaning theory must be selected on the basis of "evidence plausibly available to an interpreter," that is, "someone who does not already know how to interpret utterances the theory is designed to cover" (1984b: 128).

Davidson boldly claims that nothing can be a language unless a correct compositional meaning theory that issues in a true and interpretive sentence like (S) for each sentence of that language can be selected on the basis of the sorts of observations plausibly available to a radical interpreter. In a bit more detail, a radical interpreter, by definition, is ignorant of the language she is trying to interpret, and also lacks access to bilingual informants, prior dictionaries, and the like. A radical interpreter is generally allowed to be able to determine when an informant holds a sentence true, even though she fails to understand whatever sentence is being held true (1984b: 135). So, in effect, the primary data for radical interpretation are formulable by what might be called *singular held true sentences*, for example, (E):

(E)   Kurt belongs to the German speech community and Kurt holds true "Es regnet" on Saturday at noon and it is raining near Kurt on Saturday at noon.

Data like (E) can be collected from a variety of speakers across a variety of times by someone who does not already understand German, and will eventually confirm a generally held true sentence like (GE):

(GE) For any speaker of the German speech community and for any time, that speaker holds true "Es regnet" at that time if and only if it's raining near him or her at that time.

Claims like (GE) provide evidence for a radical interpreter that speakers of the German speech community *take* some form of words to express a specific truth (1984b: 135). But what licenses inferring from data like (GE) a target theorem like (R)?

(R)   "Es regnet" is true in German of a speaker s at a time t just in case it's raining near that speaker at that time.

Could all German speakers get it wrong? Davidson denies any such possibility, answering that the inference from (GE) to (R) is legitimate, because a certain *principle of charity* is legitimately presupposed (1984b: 137). According to this principle, the favored compositional meaning theory for a language L must entail sentences of form (T) such that most sentences speakers of L hold true are in fact true.

Under radical interpretation, sentences speakers hold true keep turning out to be true. This is no accident; rather, it's because radical interpretation is a special sort of project, namely, one that is constituted by this principle of charity. So, once data like (GE) are collected, we can infer its corresponding T-sentences (R) *via* this principle of charity. To boot, by virtue of securing a set of interpretive T-sentences for context sensitive sentences like "Es regnet," a truth theory automatically assigns interpretive T-sentences to context *in*sensitive sentences like "Schnee ist weiss." So, to a very rough approximation, suppose that, on the basis of what a radical interpreter collects, she devises for context sensitive sentences "Es schneit" and "Es ist weiss" interpretive T-sentences like (a) and (b):

(a)   For any speaker and time, if the speaker utters "Es schneit" at that time, then the speaker's utterance of "Es schneit" is true just in case it's snowing near him or her at that time.

(b)   For any speaker, time and object, if the speaker utters "Es ist weiss" at that time, then the speaker's utterance of "Es ist weiss" is true at that time of that object if, and only if that object is white at that time.

But on the basis of these generalizations, suppose that somehow or other a radical interpreter conjectures application conditions for sub-sentential expressions like "schnee" and "weiss," as in both (c) and (d):

(c)   For any object and time, "schnee" is true of that object at that time if, and only if, that object is snow at that time.

(d)   For any object and time, "weiss" is true of that object at that time if, and only if, that object is white at that time.

Such information we can imagine being exploited in attributing truth conditions (e) to a context *in*sensitive sentence like "Schnee est weiss."

(e)   "Schnee ist weiss" is true if, and only if, snow is white.

The upshot is that only those compositional meaning theories that entail T-sentences licensed by radical interpretation are adequate.

In the next few sections, we will explore some philosophical ramifications of embracing truth theories as the correct form for compositional meaning theories, and of embracing radical interpretation as their presumed manner of confirmation.

Adverbial modification

We discussed Davidson's views about the nature and role of actions and events above. Yet the strongest argument he advances for the existence of events and certain views about their nature derives, remarkably, not from any piece of pure metaphysical reasoning, but rather from constraints on compositional meaning theories.

Any theory of meaning for a language must embody a view of the relationship between language and reality. Davidson's conviction is that a compositional meaning theory, by providing a view about this relationship, offers substantive answers to the metaphysical questions about reality and its nature. In particular, the best compositional meaning theory for English, for example, will require positing events in order to explain what we say for sentences about actions, events, and singular causal relationships (like "Frank's pushing Bill caused him to fall").

So, consider the action/event sentence (1) and an obvious candidate for its interpretive truth condition (2):

1   John hit Bill.
2   "John hit Bill" is true if, and only if, John hit Bill.

In (2), linguistic expressions are both used and mentioned. Both the words "hit," "John," and "Bill" and the corresponding aspects of the world, the people and the action which relates them, are discussed. In this limited sense, (2) could be said "hook up" language and reality. This hook-up remains silent, though, on the nature of reality, since it tells us no more than what the English sentence (1) requires for its truth, namely, that John hit Bill. However, since an adequate compositional meaning theory must be finite (1984b: 4–15) in constructing a compositional meaning theory for, say, English (an unbounded language), structure must be read into its sentences. Consider action sentences (3)–(5):

3   John hit Bill at six.
4   John hit Bill at six in the bedroom.
5   John hit Bill at six in the bedroom with the stick.

These are but three examples of how adverbial modifiers can be added, in the form of prepositional phrases, to an English sentence without compromising its grammaticality or intelligibility. There is no obvious specifiable upper limit upon the number of modifiers English allows us to sensibly attach to these sorts of sentences – "after dark," "on his ear," "on a Tuesday," and so on. Therefore, any compositional meaning that

treats each of such sentences as involving a distinct primitive relation threatens to offend against the finitude condition on a compositional meaning theory. Were we to try to devise a compositional meaning theory for English according to which (3) is true just in case the three-place relation of hitting obtains between John and Bill and six o'clock, what would we say about (4)? Is it true just in case the distinct four-place relation of hitting obtains among two people, a time and a place? And on it goes. This strategy for devising interpretive T-sentences, in effect, winds up treating each adverbial modifier as introducing a distinct and novel relation with a distinct number of *relata*. A compositional meaning theory prohibits positing indefinitely many distinct primitive predicates in a language. Recall, the aim of a compositional meaning theory is to explain how there can be indefinitely many non-synonymous meaning sentences given a finite basis of meaningful components.

The method of truth in metaphysics

On the basis of such considerations, Davidson advances a proposal which simultaneously reveals common elements in all these many distinct sentences, issues in their correct interpretive T-sentences, and validates logical implications among them, for example, that (4) logically implies (3), and that both (4) and (3) imply (1). His idea, roughly, is to interpret sentences like (1) and (3)–(5) in such a way that they are "revealed" to harbor an existential quantifier that ranges over events.

The metaphysical punch line is that events must exist, since, otherwise, a *finite* compositional meaning theory for English would be unattainable. (This is how Davidson justifies the sort of ontological commitment to events we saw at work in the section "Events and Causation.") More specifically, Davidson argues that the best compositional meaning theory for English will interpret action sentences like (1) and (3)–(4) roughly along the following lines of (1′) and (3′)–(4′):

1′   *There is an event* that is a hitting of Bill by John.
3′   *There is an event* that is a hitting of Bill by John *and* it occurs at six.
4′   *There is an event* that is a hitting of Bill by John *and* it occurs at six *and* it occurs in the bedroom.

In each interpretation, there is existential quantification over an event. Each invokes exactly the same three-place relation of hitting and thereby shows what we already intuitively recognize to be common ground among them. And in (3′) and (4′), adverbial modification is transformed into a predication of this posited event.

This strategy for discerning ontological commitments extends to any locution where quantification and predication are required in order to construct a satisfactory compositional meaning theory for natural language, and thus provides us with a general method for isolating ontological commitments for speakers. Note how far removed this style of argument is from a long-standing tradition of treating metaphysics as an independent discipline, which somehow abstracts from questions about meaning or about science, instead taking a middle ground between science and analysis for discerning the nature of reality. Davidson, perhaps, could be interpreted as challenging the need for any such middle ground.

Against facts

Adopting the method of truth in metaphysics as a strategy for discerning ontological commitment also has negative ramifications for traditional ontological commitment to facts. A view going back at least to early Russell and Wittgenstein of the *Tractatus* is that the world is populated with *facts* and that true sentences are true because they *correspond* to these facts. In more recent writings, for example, "The Myth of the Subjective" (1989b) and "The Structure and Content of Truth" (1990), Davidson attempts to refute the claim that sentences are representations of reality. In the latter essay, he argues against "the popular assumption that sentences, or their spoken tokens, or sentence-like entities or configurations in our brains can properly be called 'representations', since there is nothing for them to represent" (Davidson 1990). In particular, there are *no* facts. And if there are no facts, and so facts fail to make sentences true, then we can ask, In what sense are sentences representational?

On Davidson's approach to meaning, the best compositional meaning theory does not require treating sentences as corresponding to facts or propositions. (Look in (T) above, the candidate interpretation of "Schnee ist weiss." On its right hand side, there is an English sentence, which provides an interpretation of the German sentence without making any reference to a fact or proposition.) Though not all of what Davidson says in his attack on representations can be addressed here, enough can be said about his attack against facts and correspondence theories of truth to illuminate his methodological approach.

Davidson's main argument against facts and correspondence theories of truth appears in his "True to the Facts" (in 1984b). There we find his so-called Great Fact argument for the conclusion that, given certain plausible assumptions, there can be at most one fact. The assumptions Davidson presumes are "that a true sentence cannot be made to correspond to something quite different by the substitution of co-referring singular terms, or by the substitution of logically equivalent sentences." From these assumptions he argues that one can prove that "if true sentences correspond to anything, they all correspond to the same thing" (1990: 303). Rightly or wrongly, Davidson takes these assumptions to embody traditional wisdom about facts.

The main point of the Great Fact argument is that if a context satisfies these two assumptions, then that context is truth functional (that is, the context looks just to the truth-value of the elements it relates, rather than to richer features, such as their meaning). So, if one sentence is made true by some fact *f*, then every sentence that agrees in truth-value with it is also made true by *f*, and thus, the Great Fact.

Since Davidson places the concept of truth to the forefront of the project of giving a theory of meaning for a natural language, and since he rejects one traditional correspondence theory of truth, it's appropriate to ask what his view is about this central concept. In the next section, we will consider briefly Davidson's position.

Truth and correspondence

As noted in the last section, Davidson rejects the traditional correspondence theory of truth, correspondence with facts. Still, in a way, Davidson himself is a correspondence theorist, since he explains the truth of sentences in terms of a relation between lan-

guage and something else. Recasting a correspondence theory of truth this broadly, Davidson's claim is that the sort of compositional meaning theory for a natural language he endorses is a correspondence theory of truth (1984b: 70), because it explains what it is for sentences to be true, not by relating sentences to objects, but by relating predicates and referring terms to objects *via* the relations of satisfaction and reference, and exhibiting the conditions under which sentences are true in terms of those relations.

Recall from our earlier discussion of adverbial modification, that according to Davidson, the best compositional account for explaining the unboundedness of modification as in "John kissed Mary in the park after dark" will treat this sentence as *quantifying* over an event and predicating of this event that it is a kissing by John of Mary, and that it took place in the park and that it occurred after dark. If Davidson is right, then in order for this sentence to be true, there must be an entity in the world, an event, that, so to speak, makes this sentence true.

This way of thinking about correspondence promises to be more illuminating than the traditional correspondence theory, because we can exhibit by way of the proof of a T-sentence from the axioms of a truth theory how the truth conditions of the sentence are arrived at on the basis of the satisfaction conditions for its significant parts; for each non-synonymous sentence, there will be a different route to its truth conditions. Moreover, it is clear that this approach affords no way of eliminating the semantic concept(s) on which it is based, that is, satisfaction and reference, so this approach is not a redundancy theory. (The other side of this coin is that the approach sheds little light on what it is for a predicate to be satisfied by an object or for a singular term to make reference to an object.) It is in this sense, and only in this sense, that Davidson is a correspondence theorist. Davidson, in more recent work, rejects this label as misleading (1990), though not the view itself. It is worth noting that being a correspondence theorist in this sense is neutral with respect to traditional disputes between various brands of realism and anti-realism.

Thus far we have been exploring ramifications of Davidson's views about meaning and interpretation for the nature of an external reality; in the next few sections the discussion turns inwards to explore conclusions Davidson draws about an inner reality, about the nature of mind.

Animal thought

What comes first, language or thought? Intuition cuts both ways: it's hard to find a pet owner who believes that her pet is a thoughtless brute, regardless of whether that pet has any facility, to speak of, with language. If intuition has got this right, then language and thought are independent. Yet it seems almost equally intuitive, that at least when *we* think, we think in our native tongue. So at least for we who have language it seems as if the two are mutually dependent.

Davidson, however, defends the unrestricted thesis that *any* capacity to think requires facility with language, and so only creatures with a language can think. So much for pet lovers! In "Thought and Talk" (1984b) and "Rational Animals" (1986c), he begins his argument for his controversial thesis by noting that ascription of psychological states to others, e.g., belief, desire, intention, and the like, exhibit *semantic*

opacity. So, in attributing to a dog the belief that the cat went up the tree, would the propriety of this ascription be affected were we to substitute for "the tree" another expression that refers to the same tree, say, "the chestnut in your backyard"? Given that the tree *is* the chestnut in your backyard, would you be so inclined to ascribe to the dog that he believes that the cat went up the chestnut in your backyard? If not, this would disclose that your attribution of belief to the dog falls short of literalness; that is, it is semantically opaque (1986c: 475).

Davidson's main contention is that semantic opacity, a failure to preserve truth under co-referential substitution, exists only when language is tied to thought. He advances two different lines of argument for his conclusion. The first appeals to a *holistic* thesis about belief ascription, by which is meant that we could never have grounds for ascribing a single belief to an organism except against the background of a wide array of other beliefs. Since, as Davidson argues, we could never have grounds for ascribing the required array of background beliefs to creatures that did not have a language, we could never be warranted in ascribing to such creatures any thought at all. The argument runs as follows:

1 Belief ascription exhibits semantic opacity, and
2 semantic opacity requires that we regard beliefs as possessing some definite intentional content, and
3 the possession of a belief with a definite intentional content presupposes "endless" further beliefs (holism).
4 Therefore, a creature to whom we are warranted in ascribing a belief is one that must possess a sophisticated behavioral repertoire;
5 but only linguistic behavior exhibits the sort of complex pattern that might warrant such ascription.

This argument, even if sound, at most establishes that we are unlikely ever to have decisive *evidence* that a speechless creature has beliefs. But the views for which we can collect decisive evidences and what we can establish as truth can come apart. Davidson wants to draw a stronger conclusion, that "unless there is actually such a complex pattern of behavior, there is no thought" (1986c: 476). He does so by arguing that:

6 Propositional attitudes require a dense network of beliefs (holism), and
7 "in order to have a belief, it is necessary to have the concept of belief" and
8 "in order to have the concept of belief one must have language," that is, one must be a member of a "speech community." (1986c: 478)

The big question for Davidson is how to get from the ubiquity of belief (6) and the view that beliefs require second-order beliefs (7) to his conclusion that "a creature must be a member of a speech community if it is to have the concept of belief" (8) (1984b: 170). He argues as follows:

9 The possibility of belief or thought generally is taken to depend on the concept of a representation that might be true or false, and
10 a concept of truth and falsity includes some notion of an objective, public domain, and
11 this, in turn, is possible only for an interpreter. (in 1984b: 170, 1986c: 480)

Davidson holds that only utterances can afford the fine-grained structure required for attributing thought; for only a creature whose behavior exhibits the kind of structure implied by a compositional meaning theory is a creature in which semantically opaque representations can make an appearance.

Alternative conceptual schemes

Even if thought requires language, isn't it possible that different people, communities, cultures, or periods view, conceptualize, or make the world (or their worlds) in different ways? Couldn't another thinker have concepts or beliefs radically different from our own? Davidson, unsurprisingly, identifies conceptual schemes with sets of intertranslatable languages (1984b: 185). By so doing, he transforms the question about alternative conceptual schemes into one about whether there could be non-intertranslatable languages. But why should anyone believe that questions about conceptual relativity have anything to do with translation?

Davidson's identification requires two assumptions: first, that speakers alone have thoughts and second, that any concept a speaker possesses and any thought he can entertain is expressible in his language. (Both assumptions were evident in our discussion above of "Thought and Talk," and "Rational Animals.") Together these entail that a difference in the conceptual schemes of two people requires that a portion of the language that one speaks is not translatable into any portion of the other's.

Once conceptual schemes are identified with sets of intertranslatable languages, the question of whether sense can be made of radically different conceptual schemes reduces to whether sense can be made of two non-intertranslatable languages (or, much the same, to whether or not a "significant range of sentences in one language could be translated into the other" (1984b: 185)). Davidson argues that making sense of such talk requires a criterion for when a form of behavior can be counted both as speech behavior and as speech that is untranslatable into our own. He then argues that no sense can be made of a total failure of translatability between languages (1984b: 185), and so no one could be in a position to judge that others had concepts or beliefs radically different from his own (p. 197).

In short, pressures from the nature of radical interpretation together with the fact that "all understanding of the speech of another involves radical interpretation" (1984b: 125) force him to draw his critical conclusions.

Anti-skepticism

Another consequence of radical interpretation is anti-skepticism, that is, the impossibility of massive error. In a number of articles, beginning with "The Method of Truth in Metaphysics," and including "Empirical Content," and "A Coherence Theory of Truth and Knowledge," continuing more recently through "What is Present to the Mind," "The Conditions of Thought," "Three Varieties of Knowledge," and "Epistemology Externalized," Davidson argues, on the basis of his principle of charity, that an interpreter cannot find speakers to have largely false beliefs, even if she herself has no opinion as to the general truth and falsity of these beliefs. Given what beliefs are, and how their contents are determined on this story, Davidson is committed to the

impossibility that "all our beliefs about the world might be false" (1991b: 193). A radical interpreter must have beliefs about the world in order to succeed in ascribing to others beliefs about the world. But, as radical interpretation is conceived, she also must find others largely in agreement with her in those beliefs.

Davidson's anti-skeptical argument from radical interpretation rests on two assumptions, namely, that to be a speaker is to be interpretable by others, and that to be interpretable by others requires being largely right, not only in one's general beliefs, but in beliefs about the local environment. On the assumption that radical interpretation is possible, the proper way to state the requirement on a speaker is that her beliefs about her environment be mostly true. The crucial aspect of radical interpretation is the importance of *causality* in determining what someone means or believes. We cannot "in general fix what someone means independently of what he believes and independently of what caused the belief . . . The causality plays an indispensable role in determining the content of what we say and believe" (1986: 435). So, it is the central role of causation in fixing the contents of beliefs that ensures that the truth of everything we believe is not in general "logically independent" of having those beliefs; and that others cannot differ too much from us in what they believe.

The central claim throughout is that how the contents of beliefs are determined limits the extent of falsity and diversity discriminable in a coherent set of beliefs. In interpreting another, a radical interpreter ventures hypotheses as to what in the circumstances in question causes a speaker to hold true the sentence in question. This is supposed to provide him (*ceteris paribus*) with the meaning of that sentence. In every case there will be different causal chains leading to the same utterance. A radical interpreter must choose one. He does so by responding to something in the environment, and so converges on something that is a common cause both of his own response and of the utterance of the speaker, thereby correlating the two and thus giving the content of the speaker's utterances. (This is what Davidson calls "triangulation" (1986c: 480, 1991b: 159–60). The central thought that emerges is that without constraints on what a creature is thinking about in addition to those provided simply by treating it as a rational creature capable of thought and speech, any answer to the question what it is thinking about will be so wildly underdetermined that we can give no clear content to the idea that it is thinking anything at all. Is the dog thinking that the squirrel ran up the tree or instead that large brown thing in front of it or the tree that Uncle Bill planted ten years ago, and so on. Without any clear way to rule out one of these attributions over the others, why assume that the dog believes any one of them?)

The method of radical interpretation "enforces" on any successful interpreter the conclusion that a speaker's beliefs are largely true and largely like her own. Thus, global skepticism is ruled out.

Anti-Cartesianism and first person authority

A central feature of the Cartesian tradition in modern philosophy is that at the foundation of the structure of our justified beliefs about the world are our beliefs about our own mental states, our attitudes, experiences, and sensations. As we have seen, Davidson's approach both to meaning and interpretation, and to central issues in epistemology, is *anti-Cartesian* inasmuch as he rejects this assumption. A radical interpreter

is restricted to behavioral evidence in interpreting another. From this standpoint, Davidson treats the central concepts employed in interpreting another as theoretical concepts introduced to keep track of behavior. Viewed from his perspective, the role of a theory of interpretation is to identify and systematize patterns in the behavior of speakers in relation to their environment. If this is right, we do not first have access to facts about speakers' meanings and attitudes, including our own.

How he aims to reconcile his treatment of the central concepts of interpretation as theoretical with the presumption of first person authority, that is, with the fact that speakers are necessarily more authoritative in general about their own attitudes and sensations than others are, is a central topic in Davidson's later writings.

It seems scarcely intelligible that another could be as well placed as you are with respect to whether you believe that you are hungry, or in pain. This asymmetry in epistemic position is connected with a difference in the way we know our own mental states and the way others know them. In ascribing mental states to others, we rely on their behavior (or records of their behavior); but in the case of first person ascription, we do *not*. Indeed, in our own case we do not rely on evidence at all, and so do not consult our behavior. Although knowing something in a different way or not on the basis of evidence does not in itself guarantee that what is known is known better, we may expect that this difference in how one knows one's own mental states (first person knowledge) and how others do underlies first person authority.

The challenge that first person authority presents to Davidson's assumption of the theoretical character of the concepts of interpretation is first taken up in his "First Person Authority," where he offers an explanation for the presumption that "a speaker is right when he sincerely attributes a belief, desire, or intention to his present self" (1984a: 101) by grounding it in the assumptions that an interpreter must make in order to succeed at interpretation.

Davidson aims to explain the asymmetry between our knowledge of our mental states and our knowledge of the mental states of others (or, alternatively, the asymmetry between our own knowledge of our mental states and the knowledge others have of them) by explaining a closely related asymmetry: why there is a "presumption that a speaker is right when he sincerely attributes a belief, desire, or intention to his present self, while there is no such presumption when others make similar attributions to him" (1984a: 101). His explanation of first person authority rests on an explanation of an asymmetry between the knowledge a speaker and interpreter have of meanings of the speaker's words. This asymmetry between the knowledge one has of one's own words and an interpreter's knowledge of the meanings of one's words is most striking in the case of an interpreter who is not a member of one's speech community.

The form of Davidson's argument is the following: one speaks a language only if one is interpretable; one is interpretable only if one is mostly right about the meanings of one's words; therefore, one speaks a language only if one is mostly right about the meanings of one's words.

Davidson's central methodological assumption is that a third person point of view on others' utterances and psychological states is primary in the sense that behavioral evidence forms our only evidence for the application of linguistic and psychological concepts and terms to others, and that their content is to be understood wholly in terms

of their role in accounting for the behavioral evidence available to us from this standpoint. His shift of viewpoint is so fundamental that, once adopted, the whole landscape in the philosophy of language and mind looks different. If Davidson is right, the central mistake of our philosophical tradition is the assumption of the Cartesian standpoint, and, in particular, the central place our tradition accords to the epistemic priority of knowledge of our mental states to knowledge of the world and other minds. Once this assumption is relinquished, each domain in which we have knowledge will be seen as necessary for the others, but knowledge of the world and by extension of other minds will turn out to be autonomous from knowledge of our own minds, in the sense that it is not explicable by appeal to inferences from a basis in knowledge of our own minds. In light of the alternative, Davidson's picture is attractive. Part of its interest and power lies in its promise to lay to rest what have been perhaps the central problems of the tradition from the beginning of the modern period. Despite the difficulties it faces, it is worth pursuing.

The rejection of empiricism

Another consequence of his taking what we might call a third person perspective of the radical interpreter as methodologically fundamental is the rejection of all forms of traditional empiricism. Essential to traditional empiricism is its attempt to account for our knowledge of the world exclusively by appeal to sensory experience. What is distinctive about empiricism is not the thought that sensory experience can play a role in justifying our beliefs about the world around us, but that it plays the role of a foundation for our empirical knowledge. This in turn entails that the first person point of view is fundamental, since each person's experience is treated as being his own foundation for his empirical knowledge. In adopting the third person point of view as fundamental, then, Davidson rejects a central tenet of all forms of empiricism, and the traditional project associated with it of explaining our empirical knowledge by appeal to experience. Rather, in Davidson's view, our knowledge of the world around us, of other minds, and of our own minds, has a unified source in our nature as rational beings capable of communicating with one another.

In conclusion, Davidson argues that language, mind, and action are inseparable. To account for language, that is, to answer the question, What is meaning?, he advances the radical idea that a theory of meaning can be satisfactory only if it discovers a finite basic vocabulary and rules of composition in the language to be interpreted. The aim to provide a comprehensive understanding of natural languages led him to a treatment of the theory of truth for a language as an empirical theory, and to the adoption of the stance of the radical interpreter as the standpoint for confirmation, linking the structure of a rich theory with its basic evidence, and placing the theory of meaning in the context of a theory of rational agency. Adopting this stance as fundamental is tantamount to the rejection of Cartesianism and empiricism, and so the abandonment, among other philosophical mainstays, of conceptual relativism, global skepticism, and representationalism. Theories frequently yield insight into problems that they were not specifically designed to solve. As with other significant philosophers, a careful reading of Davidson's writings bears out both how broad in scope his philosophical accomplishments are and, more importantly, how well they cohere.

Note

Some of this entry, in particular, from the section "Against Facts" to the end, is adapted from Kirk Ludwig's and my forthcoming *Donald Davidson: Truth, Meaning, and Reality*, Oxford University Press.

Bibliography

Works by Davidson

1980: *Essays on Actions and Events*, Oxford: Oxford University Press.
  The collection contains the following essays, listed here with their original publication dates, and page extents in the collection.
  "Actions, Reasons, and Causes" [1963], pp. 3–19.
  "Causal Relations" [1967], pp. 149–62.
  "Mental Events" [1970], pp. 207–27.
  "Psychology as Philosophy" [1973], pp. 229–44.
  "The Individuation of Events" [1969], pp. 163–80.
  "The Logical Form of Action Sentences" [1967], pp. 105–48.
  "The Material Mind" [1973], pp. 245–59.
1984a: "First Person Authority," *Dialectica* 38, pp. 101–11.
1984b: *Inquiries into Truth and Interpretation*, Oxford: Oxford University Press.
  The collection contains the following essays, listed here with their original publication dates and page extents in the collection.
  "In Defence of Convention T" [1970], pp. 65–76.
  "On the Very Idea of a Conceptual Scheme" [1974], pp. 183–98.
  "Radical Interpretation" [1973], pp. 125–39.
  "The Method of Truth in Metaphysics" [1977], pp. 199–214.
  "Theories of Meaning and Learnable Languages" [1965], pp. 3–15.
  "Thought and Talk" [1975], pp. 155–70.
  "True to the Facts" [1969], pp. 37–54.
  "Truth and Meaning" [1967], pp. 17–36.
1986a: "A Coherence Theory of Truth and Knowledge" [1983], in LePore 1986, pp. 307–19.
1986b: "Empirical Content" [1982], in LePore and McLaughlin 1986, pp. 320–32.
1986c: "Rational Animals" [1982], in LePore and McLaughlin 1986, pp. 473–80.
1989a: "Conditions on Thought," in *The Mind of Donald Davidson*, issue 36 of *Grazer Philosophische Studien* (ed. Brandl and Gombocz), Amsterdam: Rodopi, pp. 193–200.
1989b: "The Myth of the Subjective," in *Relativism: Interpretation and Confrontation*, ed. M. Karusz, South Bend, IN: University of Notre Dame Press.
1989c: "What is Present to the Mind," in *The Mind of Donald Davidson*, issue 36 of *Grazer Philosophische Studien* (ed. Brandl and Gombocz), Amsterdam: Rodopi, pp. 3–18.
1990: "The Structure and Content of Truth," *Journal of Philosophy* 87, pp. 279–328.
1991a: "Epistemology Externalized," *Dialectica* 45, pp. 191–202.
1991b: "Three Varieties of Knowledge," in *A. J. Ayer: Memorial Essays*, ed. A. Phillips, Cambridge: Cambridge University Press, pp. 153–66.

Works by other authors

Fodor, J. and LePore, E. (1992) *Holism: A Shopper's Guide*, Oxford: Blackwell Publishers.
Hahn, L. (ed.) (2000) *The Philosophy of Donald Davidson*, La Salle, IL: Open Court.

LePore, E. (ed.) (1986) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford: Blackwell Publishers.

LePore, E. and McLaughlin, B. (eds.) (1986) *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Oxford: Blackwell Publishers.