Introduction

KIRK LUDWIG

Donald Davidson has been one of the most influential philosophers working in the analytic tradition in the last half of the twentieth century. He has made seminal contributions to a wide range of subjects: the philosophy of language and the theory of meaning, the philosophy of action, the philosophy of mind, epistemology, metaphysics, and the theory of rationality. His principal work, spread out in a series of articles stretching over nearly forty years, exhibits a unity rare among philosophers contributing to so many different topics. His essays are elegant, but they are also noted for their compact, sometimes cryptic style, and for their difficulty. Themes and arguments in different essays overlap, and later papers often presuppose familiarity with earlier work. Together, they form a mosaic that presents a systematic account of the nature of human thought, action, and speech, and their relation to the natural world, that is one of the most subtle and impressive systems to emerge in analytic philosophy in the last fifty years.

The unity of Davidson's work lies in the central role that reflection on how we are able to interpret the speech of another plays in understanding the nature of meaning, the propositional attitudes (beliefs, desires, intentions, and so on), and our epistemic position with respect to our own minds, the minds of others, and the world around us. Davidson adopts as methodologically basic the standpoint of the interpreter of the speech of another whose evidence does not, at the outset, presuppose anything about what the speaker's words mean or any detailed knowledge of his propositional attitudes. This is the position of the radical interpreter. The adoption of this position as methodologically basic rests on the following principle:

> The semantic features of language are public features. What no one can, in the nature of the case, figure out from the totality of the relevant evidence cannot be part of meaning. (Davidson 1984a [1979], p. 235)

The point carries over to the propositional attitudes, whose attributions to speakers are inseparable from the project of interpreting their words.

Virtually all of Davidson's major contributions are either components of this project of understanding how we are able to interpret others, or flow from his account of this. Davidson's work in the philosophy of action helps to provide part of the background for the interpreter's project: for an understanding of the nature of agency and rationality is also central to understanding the nature of speech. Davidson's work on the structure of compositional meaning theories plays a central role in understanding how we can interpret others as speakers; it also contributes to an understanding of the nature of agency through applications to the logical form of action sentences and connected investigations into the nature of events. The analysis of the nature of meaning and the attitudes through consideration of radical interpretation leads in turn to many of Davidson's celebrated theses in the philosophy of language, mind, and knowledge.

This introduction briefly surveys Davidson's life and philosophical development (§§1–2), and then provides an overview of major themes in, and traces out connections between, his work on the theory of meaning (§3), the philosophy of action (§4), radical interpretation (§5), philosophical psychology (§6), epistemology (§7), the metaphysics of events (§8), the concept of truth (§9), rationality and irrationality (§10), and the theory of literature (§11). The final section provides a brief overview of the volume.

1. EARLY LIFE AND INTELLECTUAL DEVELOPMENT

Donald Davidson was born on March 6, 1917, in Springfield, Massachusetts. After early travels that included three years in the Philippines, the Davidson family settled on Staten Island in 1924. From 1926, he attended the Staten Island Academy, and then began studies at Harvard in 1935, on a scholarship from the Harvard Club of New York. During his sophomore year, Davidson attended the last seminar given by Alfred North Whitehead, on material from *Process and Reality* (Whitehead 1929). Of his term paper for the seminar, Davidson has written, "I have never, I'm happy to say, received a paper like it" (Davidson 1999a, p. 14; henceforth parenthetical page numbers refer to this essay). He received an 'A+'. Partly inspired by this experience, as an undergraduate Davidson thought that in philosophy "[t]ruth, or even serious argument, was irrelevant" (p. 14).

For his first two years at Harvard, he was an English major, but he then turned to classics and comparative literature. His undergraduate education in philosophy, aside from his contact with Whitehead, came through a tutor

in philosophy, David Prall, and from preparing for four comprehensive exams – in ethics, history of philosophy, logic, and metaphysics. His main interests in philosophy at the time were in its history and its relation to the history of ideas.

He graduated in 1939. That summer he was offered a fellowship at Harvard in classical philosophy. He took his first course in logic with W. V. Quine, on material from *Mathematical Logic* (Quine 1940), which was published that term. Davidson's fellow graduate students at Harvard included Roderick Chisholm, Roderick Furth, Arthur Smullyan, and Henry Aiken.

Quine changed Davidson's attitude toward philosophy. Davidson reports that he met Quine on the steps of Eliot Hall after interviewing as a candidate to become a junior fellow. When Quine asked him how it had gone, Davidson "blurted out" his views on the relativity of truth to a conceptual scheme. Quine asked him (presciently borrowing an example of Tarski's) whether he thought that 'Snow is white' is true iff snow is white. Davidson writes: "I saw the point" (p. 22). In his first year as a graduate student, he took a seminar of Quine's on logical positivism: "What mattered to me," Davidson reported, "was not so much Quine's conclusions (I assumed he was right) as the realization that it was possible to be serious about getting things right in philosophy" (p. 23).[1]

With the advent of the Second World War, Davidson joined the navy, serving as an instructor on airplane spotting. Discharged in 1945, he returned to Harvard in 1946, and the following year took up a teaching position at Queens College, New York. (Carl Hempel was a colleague, whom Davidson later rejoined at Princeton; Nicholas Rescher was a student in one of Davidson's logic courses during this period.) On a grant from the Rockefeller Foundation for the 1947–48 academic year, Davidson finished his dissertation on Plato's *Philebus* (published eventually in 1990 [Davidson 1990b]) in Southern California, receiving the Ph.D. from Harvard in 1949.

In January 1951, Davidson left Queens College to join the faculty at Stanford, where he taught for sixteen years before leaving for Princeton in 1967. Davidson taught a wide range of courses at Stanford, reflecting his interests in nearly all areas of philosophy: logic, ethics, ancient and modern philosophy, epistemology, philosophy of science, philosophy of language, music theory, and ideas in literature, among others.

Through working with J. J. C. McKinsey and Patrick Suppes at Stanford, Davidson became interested in decision theory, the formal theory of choice behavior. He discovered a technique for identifying through choice behavior an agent's subjective utilities (the values agents assign to outcomes) and subjective probabilities (the degree of confidence they have

that an outcome will occur given an action), only to find later that Ramsey had anticipated him in 1926. This led to experimental testing of decision theory with Suppes, the results of which were published in *Decision Making: An Experimental Approach* (Davidson and Suppes 1957).

This early work in decision theory had an important influence on Davidson's later work in the philosophy of language, especially his work on radical interpretation. Davidson drew two lessons from it. The first was that in "putting formal conditions on simple concepts and their relations to one another, a powerful structure could be defined"; the second was that the formal theory itself "says nothing about the world," and that its content is given in its interpretation by the data to which it is applied (p. 32). This theme is sounded frequently in Davidson's essays.[2] The first suggests a strategy for illuminating a family of concepts too basic to admit of illuminating analyses individually. The second shows that the illumination is to be sought in the empirical application of such a structure.

At this time, Davidson also began serious work on semantics, prompted by the task of writing an essay on Carnap's method of extension and intension for the Library of Living Philosophers volume on Carnap (Davidson 1963), which had fallen to him after the death of McKinsey, with whom it was to have been a joint paper. Carnap's method of intension and extension followed Frege in assigning to predicates both intensions (meanings) and extensions (sets of things predicates are true of). In the course of work on the project, Davidson became seminally interested in the problem of the semantics of indirect discourse and belief sentences. Carnap, following Frege, treated the 'that'-clause in a sentence such as 'Galileo said that the Earth moves' as referring to an intension – roughly, the usual meaning of 'the Earth moves'. For in these "opaque" contexts, expressions cannot be intersubstituted freely merely on the basis of shared reference, extension, or truth value. Davidson became suspicious, however, of the idea that in opaque contexts expressions refer to their usual intensions, writing later that "[i]f we could recover our pre-Fregean semantic innocence, I think it would seem to us plainly incredible that the words 'The earth moves', uttered after the words 'Galileo said that', mean anything different, or refer to anything else, than is their wont when they come in other environments" (Davidson 1984 [1968], p. 108).

The work on Carnap led Davidson serendipitously to Alfred Tarski's work on truth. At Berkeley, Davidson presented a paper on his work on Carnap; the presentation was attended by Tarski. Afterward, Tarski gave him a reprint of "The Semantic Conception of Truth and Foundations of Semantics" (Tarski 1944). This led to Tarski's more technical "The Concept

of Truth in Formalized Languages" (Tarski 1983 [1932]). Tarski shows there how to provide a recursive definition of a truth predicate for a formal language that enables one to say for each sentence of the language, characterized in terms of how it is built up from its significant parts, under what conditions it falls in the extension of the truth predicate. Tarski's work struck Davidson as providing an answer to a question that had puzzled him, a question concerning accounts of the semantic form of indirect discourse and belief sentences: how does one tell when a proposed account is correct? The answer was that it was correct if it could be incorporated into a truth definition for the language in roughly the style outlined by Tarski. For this would tell one, in the context of a theory for the language as a whole, what contribution each expression in each sentence in the language makes to fixing its truth conditions. Moreover, such a theory makes clear how a finite being can encompass a capacity for understanding an infinity of nonsynonymous sentences. These insights were the genesis of two foundational papers in Davidson's work on natural language semantics, "Theories of Meaning and Learnable Languages" (Davidson 1984 [1966]) and "Truth and Meaning" (Davidson 1984 [1967]). In the former, Davidson proposed as a criterion for the adequacy of an analysis of the logical form of a sentence or complex expression in a natural language that it not make it impossible for a finite being to learn the language of which it was a part. In the latter, he proposed that a Tarski-style truth theory, modified for a natural language, could serve the purpose of a meaning theory for the language, without appeal to meanings, intensions, or the like.

Another important influence on Davidson during his years at Stanford was Michael Dummett, who lectured a number of times at Stanford during the 1950s on Frege and the philosophy of language.

During the 1958–59 academic year, Quine was a fellow at the Center for Advanced Study in the Behavioral Sciences at Stanford, where he put the finishing touches on the manuscript of *Word and Object* (Quine 1960). Davidson, who was on a fellowship from the American Council of Learned Societies that year, accepted Quine's invitation to read the manuscript. Quine's casting, in *Word and Object*, of the task of understanding linguistic communication in the form of an examination of the task of radical translation had a tremendous impact on Davidson. The radical translator must construct a translation manual for another's language solely on the basis of a speaker's dispositions to verbal behavior, without any antecedent knowledge of his thoughts or what his words mean. The central idea, that there can be no more to meaning than can be gleaned from observing a speaker's behavior, is a leitmotif of Davidson's philosophy of language. The

project of radical interpretation, which assumes a central role in Davidson's philosophy, is a direct descendant of the project of radical translation.[3] As we will see, Davidson brings together in this project the influence of both Tarski and Quine.

While at Stanford, Davidson also became interested in general issues in the philosophy of action, in part through his student Dan Bennett, who spent a year at Oxford and wrote a dissertation on action theory inspired by the discussions then going on at Oxford. The orthodoxy at the time was heavily influenced by Wittgenstein's *Philosophical Investigations* (Wittgenstein 1950). It held that explaining an action by citing an agent's reasons for it was a matter of *redescribing* the action in a way that placed it in a larger social, linguistic, economic, or evaluative pattern, and that, in particular, action explanation was *not* a species of causal explanation, which was taken to be, in A. I. Melden's words, "wholly irrelevant to the understanding" of human action (Melden 1961, p. 184). Davidson famously argued, against the orthodoxy, in "Actions, Reasons, and Causes" (Davidson 1980 [1963]), that action explanations *are* causal explanations, and so influentially as to establish this position as the new orthodoxy.

This interest in action theory connects in a straightforward way with Davidson's work on decision theory. Davidson's work on semantics and action theory came together in his account of the logical form of action sentences containing adverbial modification. Additionally, Davidson's work on action theory and decision theory, as noted earlier, provides part of the background and framework for his work on radical interpretation.

Davidson's first ten years at Stanford were a period of intense intellectual development, though accompanied by relatively few publications. During the 1960s, Davidson published a number of papers that changed the philosophical landscape and immediately established him as a major figure in analytic philosophy. Principal among these were "Actions, Reasons, and Causes" (Davidson 1980 [1963]), "Theories of Meaning and Learnable Languages" (Davidson 1984 [1966]), "Truth and Meaning" (Davidson 1984 [1967]), "The Logical Form of Action Sentences" (Davidson 1980b [1967]), "Causal Relations" (Davidson 1980a [1967]), "On Saying That" (Davidson 1984 [1968]), "True to the Facts" (Davidson 1984 [1969]), and "The Individuation of Events" (Davidson 1980 [1969]). (Details of these contributions are discussed below.) In 1970, Davidson gave the prestigious John Locke Lectures at Oxford University on the topic, "The Structure of Truth."

Davidson taught at Princeton from 1967 to 1970, serving as chair of the Philosophy Department for the 1968–69 academic year. He was appointed professor at the Rockefeller University in New York in 1970; he

moved to the University of Chicago as a University Professor in 1976, when the philosophy unit at Rockefeller University was disbanded. In 1981, he moved to the Philosophy Department at the University of California at Berkeley.

2. WORK CIRCA 1970 TO THE PRESENT

Davidson's work during the late 1960s and 1970s developed in a number of different directions.

(1) *Philosophy of action.* In a series of papers, Davidson continued to defend, refine, and elaborate the view of actions as bodily movements and action explanations as causal explanations originally introduced in "Actions, Reasons, and Causes." These papers included "How Is Weakness of the Will Possible?" (Davidson 1980b [1970]), "Action and Reaction" (Davidson 1970), "Agency" (Davidson 1980a [1971]), "Freedom to Act" (Davidson 1980a [1973]), "Hempel on Explaining Action" (Davidson 1980a [1976]), and "Intending" (Davidson 1980 [1978]). The work on the semantics of action sentences led to additional work on the semantics of sentences containing noun phrases referring to events – specifically, "Causal Relations" (Davidson 1980a [1967]), "The Individuation of Events" (Davidson 1980 [1969]), "Events as Particulars" (Davidson 1980a [1970]), and "Eternal vs. Ephemeral Events" (Davidson 1980b [1971]).

(2) *Philosophical psychology.* The publication in 1970 of "Mental Events" (Davidson 1980c [1970]) was a seminal event in the philosophy of mind. In it, Davidson proposed a novel form of materialism called anomalous monism. Davidson advanced an argument for a token-token identity theory of mental and physical events – according to which every particular mental event is also a particular physical event – that relied crucially on a premise that denied even the nomic reducibility of mental to physical properties. This was followed by a number of other papers elaborating on this theme, including "Psychology as Philosophy" (Davidson 1980 [1974]), "The Material Mind" (Davidson 1980b [1973]), and "Hempel on Explaining Action" (Davidson 1980a [1976]). Another paper from this period on the philosophy of psychology is "Hume's Cognitive Theory of Pride" (Davidson 1980b [1976]), which interprets Hume's theory of pride in the light of Davidson's causal theory of action explanation.

(3) *Natural language semantics.* Davidson elaborated and defended his proposal for using a Tarski-style truth theory to pursue natural language semantics in "In Defense of Convention *T*" (Davidson 1984a [1973]) and

extended a key idea (*parataxis*; see Chapter 1, §7, for a brief overview) of the treatment of indirect discourse introduced in "On Saying That" (Davidson 1984 [1968]) to quotation and to sentential moods (the indicative, imperative, and interrogative moods) in "Quotation" (Davidson 1984c [1979]) and "Moods and Performances" (Davidson 1984b [1979]), respectively. In addition, he edited, with Gilbert Harman, two important collections of essays on natural language semantics: *Semantics of Natural Language* (Davidson and Harman 1977) and *The Logic of Grammar* (Davidson and Harman 1975).

(4) *Radical interpretation.* Among the most important developments in Davidson's work in the philosophy of language during the 1970s was his elaboration of the project of radical interpretation, already adumbrated in "Truth and Meaning" (Davidson 1984 [1967]). Radical interpretation can be seen as an application of the insight – prompted by Davidson's work in decision theory during the 1950s – that a family of concepts whose members resist reduction to other terms one by one can be illuminated by examining the empirical application of the formal structure that they induce. The relation of the project of radical interpretation to understanding linguistic communication and meaning is taken up in "Belief and the Basis of Meaning" (Davidson 1984a [1974]) and, in the context of a defense of the claim that thought is not possible without a language, in "Thought and Talk" (Davidson 1984 [1975]). "Reply to Foster" (Davidson 1984 [1976]) contains important clarifications of the project and its relation to using a truth theory as a theory of interpretation; it responds to a critical paper by John Foster (Foster 1976), which appeared in an important collection of papers edited by Gareth Evans and John McDowell (Evans and McDowell 1976). "Reality without Reference" (Davidson 1984b [1977]) and "The Inscrutability of Reference" (Davidson 1984a [1979]) are applications of reflections on radical interpretation to the status of talk about the reference of singular terms and the extensions of predicates in a language. Davidson draws the startling conclusion (first drawn by Quine [1969]) that there are many different reference schemes that an interpreter can use that capture equally well the facts of the matter concerning what speakers mean by their words.

(5) *Epistemology.* "On the Very Idea of a Conceptual Scheme" (Davidson 1984b [1974]) originated in the last of Davidson's six John Locke Lectures in 1970 and was delivered in the published form as his presidential address to the Eastern Division meeting of the American Philosophical Association in 1973. An influential paper, it argues against the relativity of truth to a conceptual scheme and against the possibility of there being radically different

conceptual schemes. "The Method of Truth in Metaphysics" (Davidson 1984a [1977]) is concerned with the relation between semantic theory and the nature of reality. In it, Davidson argues for two connected theses about the relation between our thought and reality. The first is that the ontological commitments of what we say are best revealed in a theory of truth for the languages we speak. The second is that massive error about the world, including massive error in our empirical beliefs, is impossible. The second thesis rests in part on conclusions reached in reflections on the project of radical interpretation, especially reflections about the need to employ in interpretation what is called the Principle of Charity, an aspect of which is the assumption that most of a speaker's beliefs about his environment are true.

(6) *Metaphor.* The last development in Davidson's work during the 1970s is an important and original account of the way in which metaphors function. In "What Metaphors Mean" (Davidson 1984 [1978]), Davidson argued that it is a mistake to think that metaphors function by virtue of having a special kind of meaning – metaphorical meaning; instead, they function in virtue of their literal meanings to get us to see things about the world. "Metaphor makes us see one thing as another by making some literal statement that inspires or prompts the insight" (Davidson 1984 [1978], p. 261).

Two collections of Davidson's papers appeared during the 1980s – *Essays on Actions and Events* (Davidson 1980a) and *Inquiries into Truth and Interpretation* (Davidson 1984b). These works collected many of his papers, respectively, on the philosophy of action and the metaphysics of events, and in the theory of meaning and philosophy of language. In 1984, an important conference on Davidson's work (dubbed "Convention D" by Sydney Morgenbesser), which brought together more than 500 participants, was organized at Rutgers University by Ernest Lepore, out of which came two collections of papers – *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (Lepore and McLaughlin 1985) and the similarly subtitled *Truth and Interpretation* (Lepore 1986). A collection of essays on Davidson's work in the philosophy of action, with replies by Davidson, edited by Bruce Vermazen and Merrill Hintikka, *Essays on Davidson: Actions and Events* (Vermazen and Hintikka 1985), appeared in 1985.

Davidson's work during the 1980s can be divided into five main categories. (1) In the first category are those papers following up on issues in action theory – "Adverbs of Action" (Davidson 1985a) and "Problems in the Explanation of Action" (Davidson 1987b). (2) In the second are papers on the nature of rationality and irrationality – "Paradoxes of Irrationality" (Davidson 1982), "Rational Animals" (Davidson 1985 [1982]),

"Deception and Division" (Davidson 1985b), and "Incoherence and Irrationality" (Davidson 1985c). (3) The third category combines elements of work on the determination of thought content and epistemology. "Empirical Content" (Davidson 2001a [1982]), "A Coherence Theory of Truth and Knowledge" (Davidson 2001 [1983]), "Epistemology and Truth" (Davidson 2001a [1988]), "The Conditions of Thought" (Davidson 1989), "The Myth of the Subjective" (Davidson 2001b [1988]), and "What Is Present to the Mind?" (Davidson 2001 [1989]) are all concerned with the thesis that the contents of our thoughts are individuated in part by their usual causes in a way that guarantees that most of our empirical beliefs are true. "First Person Authority" (Davidson 2001 [1984]) and "Knowing One's Own Mind" (Davidson 2001 [1987]) are concerned to argue that knowledge of our own minds can be understood in a way that does not give primacy to the subjective, and that the relational individuation of thought content is no threat to our knowledge of our thoughts. (4) The fourth category of papers includes those that develop earlier work in the philosophy of language. "Toward a Unified Theory of Meaning and Action" (Davidson 1980b) explicitly combines decision theory with Davidson's earlier work on radical interpretation, and "A New Basis for Decision Theory" (Davidson 1985d) outlines a procedure for identifying logical constants by finding patterns among preferences toward the truth of sentences. In "Communication and Convention" (Davidson 1984 [1983]), Davidson takes up the question of what role convention plays in communication, and in particular the question of whether it is essential to communication at all. "Communication and Convention" already contains the main themes, if not so provocatively stated, of Davidson's later and more controversial "A Nice Derangement of Epitaphs," in which he argues that "there is no such thing as a language, not if a language is anything like what many philosophers and linguists have supposed" (Davidson 1986c, p. 446). "James Joyce and Humpty Dumpty" (Davidson 1991b) is another excursion into literary theory. (5) The fifth category is work on issues in ethical theory from the standpoint of radical interpretation, the Lindley Lectures, *Expressing Evaluations* (Davidson 1984a), and "Judging Interpersonal Interests" (Davidson 1986b), a central thesis of which is that communication requires shared values as much as shared beliefs.

In 1989, Davidson gave the John Dewey Lectures at Columbia, "The Structure and Content of Truth" (Davidson 1990d), echoing the title of the John Locke Lectures delivered almost twenty years before. These provide a comprehensive overview and synthesis of Davidson's work in the theory of meaning and radical interpretation up through the end of the 1980s.