Notre Dame Journal of Formal Logic
Volume 23, Number 2, April 1982

Meaning and Interpretation

MICHAEL ROOT and JOHN WALLACE*

Frank Ramsey called Russell's theory of descriptions a "paradigm of philosophy". Russell first presented the theory in his paper "On Denoting", and for many years the course of the philosophy of language was set by this paper. In 1951, approximately fifty years after the publication of Russell's famous paper, Quine published "Two Dogmas of Empiricism" and the philosophy of language took a radically different turn. Michael Dummett has called Quine's paper "probably the most important philosophical article written in the last half century" ([7], p. 375).

The interest in language that Russell's paper prompted was a somewhat local interest. There were some central or prominent problems of philosophy that philosophers, following Russell's lead or example, sought to solve or dissolve by a careful analysis of some word or phrase or by attention to the logic or grammar of words like 'the' and 'Scott'. The interest in language to which Quine's paper led, on the other hand, was a broader interest. Language, on Quine's view, is a single, articulated structure. The importance or significance of each expression is to be understood in terms of the contribution it makes not merely to the understanding of whole sentences but to the understanding of the entire language. The mastery of language is to be thought of as the mastery of a theory. The theory is tested by the data of linguistic behavior but not in any piecemeal way. Analytical hypotheses, rules of logic or grammar for words or phrases, for adverbs or denoting expressions, do not face the tribunal of evidence individually but only as a corporate body.

---

*The individual contributions of the authors of this paper are neither equal nor symmetrical. Root did most of the thinking and virtually all of the writing. Successive drafts, over the summer of 1981, were shaped by early morning running conversations in which Wallace put in his two cents worth.

Received September 9, 1981

Quine went on to develop his approach to language study in Chapters I and II of *Word and Object*, but Donald Davidson, more than Quine himself, has extended and elaborated Quine's ideal of an "empiricism without dogmas" and Quine's radical thought experiment into a philosophical program, a program that draws the philosophy of mind and action and the philosophy of language together into a comprehensive theory of human understanding.

1 Davidson's comprehensive program On Davidson's view, a theory of meaning is the central thread in a lace of theories that together enable a person to understand or interpret the actions of other people. Interpretation, here, means rational explanation. We explain a person's action by showing that it was the reasonable thing to do given his attitudes, his beliefs and desires. If we know what his attitudes are, then we can explain the person's action by describing the action in an especially revealing way: we describe the action as the action which he believed would bring him his fondest desire. Speech is of prime importance here. First, it is through a person's speech that we learn what his attitudes are, and, second, speech is itself an action and, consequently, itself the proper object of rational explanation.

When a person's action is an act of speaking, the act of uttering a sentence, *s*, the most revealing description is a description of what the person said in uttering the sentence on this occasion. This is a redescription of the act of uttering *s* as an act of indirect speech, as an act of saying that *p*. This redescription does not itself explain *why* the speaker performed the action, but it can make the action appear reasonable in light of our estimate of the speaker's beliefs and desires and, in this way, provide us with a rational explanation of why he said that *p* and of why he uttered *s*.

An example. We know Kurt's frame of mind. He wants to inform Max that it is raining, and he believes that the best way to inform him of this is to say to Max that it is raining. Kurt looks Max in the eye and utters the sentence 'Es regnet'. Why did Kurt make *these* noises? All that we need to know in order to answer this question is that in uttering 'Es regnet' on this occasion, Kurt said that it was raining. According to Davidson, this knowledge is part of what we would know if we knew a theory of meaning for Kurt's language. Given a theory of meaning and a description of the sentences a person speaks and the occasions on which he speaks them, an interpreter of the person's action can determine what he said on an occasion, and, having determined what he said, the interpreter can proceed to explain why he chose to say it.

There is one sentential attitude that, according to Davidson, provides a natural or especially apt focus for a theory of understanding. It is the attitude of holding a sentence true. Evidence for the holding true attitude comes from speech acts: a person probably holds a sentence true when he uses it to make an assertion; he probably holds it true when he is queried with the sentence and assents. Holding a sentence true has both a mental and a behavioral side. On the mental side, it is an attitude and a form of belief. But it is a special *de re* belief, viz., a speaker believes of a sentence that it is true, that is especially open to observation: it can be observed to hold between a person and a sentence without the observer knowing what the sentence means.

On Davidson's view, knowledge of a theory of meaning for a language *L* should enable an interpreter to explain why a speaker of *L* holds true the sentences of *L* on the occasions on which he holds them true. These explanations have two parts. One part has to do with the attitudes of the speaker who holds the sentences true, and the other part has to do with the meaning of the sentences he holds true. Like any explanation of an act of speech, an explanation of why a speaker holds a sentence true on a particular occasion combines an interpretation or redescription of the words of the speaker with a portrait of his attitudes. However, given the special nature of the action, viz., holding a sentence true, the interpreter can assume that the agent's desire is in every case the same: to speak the truth. Consequently, all the interpreter needs to do in order to explain why a speaker holds a sentence true is to say what the sentence means and what the speaker believes about the occasion on which he holds it true that makes his holding *it* true on *that* occasion the reasonable thing to do.

There are a number of theories of meaning or approaches to language study that do not lead to explanations of human action and, in particular, do not lead to rational explanations of why a person holds a sentence true. Quine's own approach is one of these. In *Word and Object*, Quine looks at language from the perspective of translation. However, as Davidson has pointed out, translation cannot be the central thread in interpretation. A translation manual is an effective set of rules for mapping sentences of one language into the sentences of another. Knowledge of a translation manual enables us to interpret the speech of another person, i.e., to say what he said, only if we are already able to interpret the language into which the speech is translated. We do not explain why Kurt holds 'Es regnet' true at time *t*, for example, by saying that Kurt believes that it is raining at *t* and 'It is raining' translates 'Es regnet'. Even if Kurt knew this fact of translation, this knowledge would not make his act of holding 'Es regnet' true the reasonable thing to do, if he did not know what either sentence means.

2 The data With a theory of meaning for a language, a person should be able to explain why sentences are held true, but, according to Davidson, it is the success of the theory in offering these explanations that confirms the theory and gives an interpreter reason to believe that the theory is true. Here Davidson adopts Quine's thought experiment of trying to penetrate the language and mind of a hitherto unknown people. What data are available to an interpreter who doesn't have the benefit of a bilingual informant or any insight into the nature of the natives' minds or language? It is on the basis of these data that the interpreter will build and test his theories about these people. The data have to be rich enough to confirm the theory but bare enough to be understood by anyone who is ignorant of the theory. Davidson maintains that data on sentences held true serve this purpose. Observations about the content of the natures' beliefs or desires, on the other hand, are not bare enough; they depend on prior knowledge of what the natives' sentences mean and on a prior ability to interpret their speech.

How do the data that support radical interpretation differ from the data that, according to Quine, support radical translation? Quine distinguishes

between standing and occasion sentences. The verdict on occasion sentences changes from occasion to occasion. The verdict on standing sentences seldom varies. Speakers hand down verdicts on sentences by assenting to them or dissenting from them when queried on given occasions. In short, occasion sentences are sentences that contain, in some manifest or latent sense, an indexical expression. As the referent of this expression changes from occasion to occasion so does the speaker's verdict to the sentence. For Quine, verdicts on occasion sentences are the primary data to which translation manuals must be true.

A similar distinction can be drawn between sentences held true. Some of the sentences that a speaker holds true are occasion sentences and other sentences are standing sentences. An occasion sentence, from the perspective of radical interpretation, is a sentence sometimes held true and sometimes false according to "discoverable changes in the world" (see [2]).

For Quine, the occasion of an occasion sentence is what prompts a speaker to assent to the sentence. For Davidson, the occasion of an occasion sentence is the occasion when a speaker holds a sentence true. For Quine, the relation between an occasion sentence and the occasion that prompts assent to the sentence is a causal relation. Quine's term is 'prompt', and the stimulation that prompts assent is a link in a chain of causes that terminates in that response. On Davidson's view, there may be no causal relation between the occasion and the act of holding the sentence true. There is only a correlation. The evidence that, according to Quine, supports radical translation is in the form:

(1) Kurt is caused to assent to 'Es regnet' at t by S.

The evidence that, according to Davidson, supports radical interpretation is in the form:

(2) Kurt holds 'Es regnet' true at t if and only if C.

From the perspective of interpretation, C is not the cause of Kurt's holding 'Es regnet' true at t, but it may be the reason why Kurt believes at t that it is raining.

What causes Kurt to assent to 'Es regnet', according to Quine, is a pattern of stimulation along Kurt's sensory surfaces. A description of this pattern replaces 'S' in (1). On the other hand, the "discoverable changes in the world" that correlate with holding sentences true or holding them false are not, always at least, changes in Kurt. A description of the occasion on which Kurt holds 'Es regnet' true, for Davidson, is a description of the weather near Kurt and not a description of Kurt's surfaces.

The descriptions of the occasions that replace 'C' in (2) are not fixed in advance. However, the descriptions are the interpreter's descriptions and are based on his conceptual scheme. It is the interpreter's conception of the occasion that he correlates with Kurt's attitude toward the sentence. However, once the interpreter develops his theory of interpretation and proceeds to interpret Kurt's speech and Kurt's attitudes, he attributes his conception of the occasion to Kurt. The interpreter observes that Kurt holds 'Es regnet' true

at t if and only if it is raining at t. On the basis of a great deal of evidence like this, the interpreter adopts a theory. On the basis of the theory, he explains his observations of Kurt's speech: Kurt believed that it was raining at t.

According to Davidson, one reason the interpreter's evidence is in the form of correlations rather than causes is that speakers make mistakes. Kurt does not always know what is going on around him; he may think it is raining when it is not or not think it's raining when it is. Consequently, there is no reason to believe that the following generalization is more than approximately true.

(3) (t)(Kurt holds 'Es regnet' true at t if and only if it is raining near Kurt at t.)

Faced with exceptions to this generalization, a radical interpreter can proceed in one of two ways. He can try to qualify the condition described on the right-hand side of the biconditional in a way that will make the statement unexceptionable. This, in effect, is what Quine's radical translater tries to purchase with his retreat to sensory stimulation. On the other hand, the interpreter can 'explain' the exceptions through the attribution of false belief. The first strategy calls on the interpreter to juggle his interpretations of Kurt's words and the second to juggle his interpretation of Kurt's attitudes. The fact that the interpreter has a decision here and that either strategy may be equally well-supported by the evidence of what Kurt holds true is the lesson, on Davidson's view, of Quine's indeterminacy thesis. That a sentence is held true is due to two factors: what the sentence means in the language of the person who holds it true and what that person believes. What the indeterminacy thesis says is that given evidence of sentences held true there is no best explanation of why a person holds them true. The interpreter can take the sentences to mean almost anything, if he makes compensatory adjustments in the attitudes he attributes to the speaker. A theory of interpretation is underdetermined by all the possible evidence of sentences held true.

Davidson's indeterminacy thesis is, in one respect at least, more radical than Quine's. According to Quine, there are some sentences whose translation is determinate. These are the observation sentences of the language, and, on Quine's view, the translation of these sentences is fixed by the conditions that prompt assent or dissent. These conditions, in the case of observation sentences, do not vary from speaker to speaker or with changes in a speaker's attitudes. In asserting or offering verdicts to these sentences, there is no opportunity for a speaker to make a mistake or for different speakers to fall into disagreement. This idea is foreign to Davidson's approach. From the perspective of radical interpretation, there are no observation sentences: no sentences that a speaker can hold true without the risk of being mistaken. The interpretation of every sentence is underdetermined by all of the available evidence.

Some of Davidson's remarks tend to cover over this point. He writes, for example:

> If interpretation is approached in the style I have been discussing, it is not likely that only one theory will be found satisfactory. The resulting indeterminacy of

> interpretation is the semantic counterpart of Quine's indeterminacy of translation. On my approach, the degree of indeterminacy will, I think, be less than Quine contemplates: this partly because I advocate adoptions of the principle of charity on an across-the-board basis . . . ([3], p. 321).

The principle of charity is the principle that other people believe pretty much as we do. This principle, according to Davidson, enables us to solve the problem of the interdependence of belief and meaning by holding belief constant while solving for meaning. This is accomplished, Davidson suggests, by assigning meaning to a person's sentences in a way that makes him right, according to our view of what is right, as often as is plausibly possible. However, given the exceptions to generalizations like (3) and the little likelihood that the interpreter will find any generalizations about Kurt and the occasions when he holds a sentence true that are exceptionless, the principle of charity leads to the implausible conclusion that the meanings of Kurt's sentences vary from occasion to occasion. If, at the outset at least, the meanings of Kurt's sentences are to be identified with the conditions under which he usually holds them true, the interpreter will have to refuse charity; when from time to time Kurt doesn't hold the sentence 'Es regnet' true when it is raining, it is reasonable for the interpreter to conclude that Kurt does not believe as he does. In short, though charity makes the interpreter reluctant to attribute mistakes to Kurt, the mere approximateness of generalizations like (3) makes these attributions unavoidable.

3 The theory Davidson maintains that a theory of truth that conforms to Tarski's Convention *T* and that has been modified to apply to a natural language can be used as a theory of interpretation. According to Tarski, a satisfactory theory of truth for a language *L* entails for each sentence *s* of *L*, a theorem of the form:

s is true-in-*L* if and only if *P*

where '*s*' is replaced by a structural-description of *s* and '*P*' by a translation of *s* into the language of the theory itself. These theorems are called *T*-sentences and a theory of truth that is satisfactory in this way is said to conform to Convention *T*. Not every theory that gives the extension of the truth predicate, i.e., entails for each sentence of *L* a true sentence of the form "*s* is true-in-*L* if and only if *P*", conforms to Convention *T*, for there are many true sentences of this form (e.g., "*s* is true-in-*L* if and only if *s* is true-in-*L*") which do not satisfy the additional formal requirements of being a *T*-sentence. The sentence used on the right of the biconditional of a *T*-sentence must be a translation of the sentence mentioned on the left.

Tarski's aim was to define the concept of an adequate account of truth by assuming the concept of translation. To test the adequacy of a Tarski-type truth theory, one only needs to inspect the theorems and see that the sentence used on the right translates the sentence mentioned on the left. Davidson's interest is radical interpretation, and here, by hypothesis, the investigator cannot recognize correct translations of the alien language into the language of his own theory. Where Tarski took translation or interpretation as basic and extracted an account of "correct account of truth", Davidson proposes

to take truth as basic and extract an account of "correct interpretation". Davidson gives us, in effect, a Convention *I* in which Tarski's direct use of the concept of translation is replaced by epistemological and evidential considerations designed to ensure indirectly that what replaces "*P*" translates what the replacement for "*s*" names. The evidence comes from observation of sentences held true. Roughly speaking, the biconditionals generated by the interpreter's theory conform to Convention *I* if most of them have the following property: the words used on the right-hand side of the biconditional describe the occasion when the speakers of the language hold the sentence on the left-hand side true.

With Tarski's theory, each theorem faces Convention *T* individually. With Davidson's "empirical" adaptation of the theory, this is no longer so; the theorems face Convention *I* and the evidence of sentences held true not individually but as a corporate body. The evidence for the interpretation of a particular sentence is evidence for the interpretation of every sentence. Davidson suggests that this empirical constraint together with certain logical constraints will ensure that a truth theory does give the interpretation of every utterance of the unknown language, for then the biconditionals that are entailed by the theory will in fact have translations of *s* replacing *P*. Such a theory would suffice for interpretation provided that the interpreter knew the evidence on behalf of the theory and knew that the evidence was adequate evidence that the theory conformed to Convention *I*. Thus, according to Davidson, a theory of interpretation in terms of truth conditions offers an informative and reasonable answer to the question of what a person could know that would enable him to understand the words of others.

If a theory of truth is to serve the purposes of interpretation, then, among other things, it should lead to an explanation of why certain sentences are held true. Does a theory of truth meet this explanation test? How would knowledge of the truth conditions of 'Es regnet' in Kurt's language enable an interpreter to explain why at time *t* Kurt held this sentence true? An interpreter who knew the truth conditions of all of the sentences of Kurt's language would know that this sentence is true when held true by Kurt at *t* if and only if it is raining near him at *t*. If this knowledge was knowledge that the interpreter had acquired empirically, in the course of theory construction, on the evidence of sentences held true, he would also know that these truth conditions interpret Kurt's words: they indicate what Kurt says when he speaks them or what they mean for him when he holds them true. If the interpreter knows that when Kurt holds 'Es regnet' true, these words mean for him that it is raining and he knows that when Kurt holds the sentence true, he believes that it is raining, then the interpreter knows that when Kurt held 'Es regnet' true, Kurt did the reasonable thing to do on that occasion. Consequently, an interpreter who knew a theory of truth for Kurt's language and knew that the theory meets the formal and empirical conditions of radical interpretation could offer a rational explanation of why Kurt holds 'Es regnet' true when it is raining.

4 Transitions If a theory is to serve the purposes of interpretation, there are purposes that it ought to serve other than the purpose of explanation and

tests that it ought to pass in addition to the explanation test. It ought to enable the interpreter to explain why sentences are held true, but knowledge of the theory ought to enable the interpreter to do other things as well. The theory should effect a series of transitions from the bare recognition of speech to an understanding of the circumstances that surround it. First, a theory of interpretation should effect a transition from the truth of an occasion sentence to a truth about the occasion when the sentence was spoken. For example, a theory of interpretation for Kurt's language should complete the following inference.

(4) 'Es regnet' when uttered by Kurt at t is true.
(5) _____
(6) It is raining near Kurt at t.

This was Frege's idea. He was interested in the information that a person could acquire by learning that a sentence was true. The information he acquires, according to Frege, is determined by the sense of the sentence and, as a result, varies according to the meaning that is attached to it. A person who knows the sense of 'Scott = the author of *Waverly*' and learns that this sentence is true is informed that Scott is the author of *Waverly*. He may be informed of more than this, but every person who knows the sense of the sentence will be informed of at least this much. This is the common thread of information that each speaker of the language will share and that makes communication between different speakers possible.

Second, a theory of interpretation should effect a transition from the falsity of a standing sentence to a truth about every occasion when the sentence is spoken. For example, if an interpreter knows the interpretation of 'Der Schnee ist schwarz' and learns that this sentence is false, he should be informed that snow is not black. A theory of interpretation for German should complete the next inference.

(7) 'Der Schnee ist schwarz' is false.
(8) _____
(9) Snow is not black.

Third, a theory of interpretation should effect transitions from sentential to propositional attitudes. One important transition is from direct to indirect discourse, i.e., from quoted speech to indirect speech when reporting another's words. Someone who knows how to interpret a person's words is able to say what the speaker said without mentioning these words. For example, if an interpreter knows the interpretation of 'Es regnet' and knows that Kurt uttered this sentence at time t, he should be informed that Kurt said that it is raining near him at t. A theory of interpretation for Kurt's language should complete this inference.

(10) Kurt uttered 'Es regnet' at time t.
(11) _____
(12) Kurt said that it is raining near him at time t.

A related and equally important transition is the transition from believing that a sentence is true to believing the proposition expressed by the sentence. A theory of interpretation should effect this transition as well.

(13) Kurt believes that 'Es regnet' is true at *t*.
(14) _____
(15) Kurt believes that it is raining near him at *t*.

This idea was important to Frege. He maintained that the sense of a sentence, i.e., what we know when we know how to understand or interpret the sentence, is the object of belief and the other attitudes. According to Frege, when a sentence forms a clause governed by a verb for a propositional attitude, the referent of the sentence is its ordinary sense. Thus we would expect that if a person was able to interpret (13), he would know whether (15) was true.

If a theory is to serve the purposes of interpretation, there are some transitions, on the other hand, that it should not effect. There is more to understanding than the interpretation of words. A theory of interpretation, for example, should not effect a transition from the interpretation of a sentence to its truth value (except perhaps for an analytic sentence), for obviously a person can understand what a sentence means without knowing whether the sentence is true. This was Frege's point in "On Sense and Reference". A theory of meaning has to allow that a person could know the meaning of 'Scott = the author of *Waverly*' and, when informed that the sentence is true, learn something new. A theory which identifies knowing the meaning of a sentence with knowing what each expression in the sentence refers to attributes too much to meaning.

In short, a theory of meaning or interpretation has to steer a course between the overly informative and the overly uninformative. Davidson maintains that the concept of truth steers it best; it effects the appropriate transitions between the recognition of speech and significant understanding. It is useful to see how this is so.

A theory of truth that conforms to Convention I will entail an infinite set of *T*-sentences, some one of which will effect the transition from the truth of each occasion sentence to a truth about the occasion when the sentence is spoken. In the example of the argument from (4) to (6), that sentence is

> 'Es regnet' is true in Kurt's language when uttered by a speaker at time *t* if and only if it is raining near the speaker at *t*.

The *T*-sentences of the truth theory effect the transition between the falsity of a standing sentence and a truth about every occasion when the sentence is spoken, e.g., from (7) to (9), in the same straightforward manner.

It is not surprising that a theory of truth should effect these transitions. The concept of truth is tailor-made for inference from word to world. But can a theory of truth effect a transition from sentential to propositional attitudes? A theory of meaning for a language should be Janus-faced. It should look from words to world, from the truth of a sentence to the conditions that contribute to that truth, and, at the same time, it should look from words to world-view, from the utterance of a sentence to the attitude or frame of mind of the speaker. If we knew a theory of meaning for Kurt's language, for example,

we should know that if (10) is true, then (12) must be true. If knowledge of a theory of truth is to enable an interpreter to rationally explain why a speaker holds a sentence true, then, as noted earlier, the interpreter must be able to redescribe the speaker's act of speech in the idiom of indirect discourse. The interpreter must describe what the speaker said in such a way that, by knowing what the speaker believes at the time of his speech and knowing that he desires to speak the truth, the interpreter knows that what he said is what he rationally ought to have said on that occasion. In short, if a theory of truth is adequate to the task of interpretation, it must effect transitions from direct to indirect speech.

In his essay "On saying that", Davidson tries to show how the meanings of sentences of indirect discourse depend upon their structure (see [1]). An analysis of these structures should reveal the relations between them. It should show, for example, that if (12) is true, then 'Kurt said something' must be true as well. Davidson wants to offer a description of the structure of these and every other sentence of the language that makes their entailments a matter of quantificational logic (see [4]). Though he does not discuss the relation between sentences like (10) and (12), we would expect that, given the analysis that Davidson offers of sentences like (12), we could show, by the rules of quantification theory, that (12) follows from (10).

Davidson proposes a paratactic semantics for sentences of indirect discourse. According to Davidson, sentences of indirect discourse are occasion sentences; they hide a demonstrative pronoun. As a result, they are true or false relative to an occasion of utterance. Moreover, when a person utters one of these sentences, he is, in fact, uttering two and not one sentence. The first sentence consists of a singular term that refers to a speaker, the two-place predicate 'said', and a demonstrative referring to the utterance of the second sentence. An utterance of the first sentence is true if and only if the utterance of the second says the same thing as (i.e., translates) some utterance of the speaker referred to by the singular term. In the case of (12), the analysis comes to this.

> 'Kurt said that it is raining near him at time t' is true in English when spoken by x at time u if and only if an utterance of 'It is raining' by x at time u translates an utterance of Kurt's at time t.

How does this analysis of (12) reveal the entailment between (10) and (12)? Davidson does not tell us how, but what he does say suggests the following. If an interpreter knew this analysis of (12) and knew that when he uttered (12) his utterance translated Kurt's utterance at t, then he would know that if (10) is true, his utterance of (12) must be true as well. The difficulty, however, is showing that someone who knew a theory of truth would know all of this. How, for example, could a person who knew only theories of truth know that his utterance of (12) translates Kurt's utterance of 'Es regnet'? In radical interpretation, the concept of translation is not given, and the interpreter has no way of knowing whether an utterance of a sentence in his own language or background theory translates an utterance of a sentence in the language of the native, Kurt.

However, what is not given might be earned. Once we have built up a theory of truth on the evidence of sentences held true, we may be in a position to introduce a concept of translation. The idea is this. Two sentences in the same language are alike in meaning if and only if there is an adequate theory of truth (a theory that meets Davidson's formal and empirical requirements) that assigns the same truth conditions to them both.

Some comment is in order. First, on this proposal, there is relativity. The idea that two sentences in a language have the same meaning becomes relative to the theory of truth we impose on the language. If there is more than one adequate theory, then a pair of sentences may be both alike and unlike in meaning. If there is slack and the evidence of sentences held true, the formal constraints and charity leave room for competing theories, then the synonymy relation, on this account, will be indeterminate. Second, to avoid circularity, the standard of sameness for truth conditions is strict. It would not do to employ an equivalence concept like that of translation here. Third, the relation we have defined is a relation between sentences and not utterances. In the case of standing sentences, the match between sentence and utterance is close, and, consequently, if the sentence 'Snow is white' translates 'Der Schnee ist weiss', then my utterance of the first sentence translates Kurt's utterance of the second. Here a translation concept for sentences can be carried smoothly into utterances. However, with occasion sentences the situation is different. The sentence 'I am angry' translates itself, but my utterance of this sentence does not translate yours. Since on Davidson's analysis of indirect discourse, these sentences refer to another's utterances, a translation concept for sentences (even if it could be drawn on the notion of truth alone) will not, in general, enable an interpreter to samesay another person. The difficulty here reflects a general point about interpreting occasion sentences: it is not enough to know the meaning of these sentences to interpret someone's utterance of them.

With two languages, A and B, the idea is somewhat more difficult. We need a theory of truth for each language, and yet we have to be able to compare the truth conditions that the two theories assign to the sentences of each. Suppose we have an adequate theory of truth for A, call it 'T_1' and an adequate theory for B, call it 'T_2'. Consider then a sentence, a, of A and a sentence, b, of B. The proposal is this. Sentence a is a translation of b if and only if the truth conditions that T_1 assigns to a = the truth conditions that T_2 assigns to b.

There is relativity here. Translation is relative to a pair of truth theories. If there is more than one adequate theory of truth for either language, then translation between some of the sentences of the two languages will be indeterminate. This may be seen as a virtue, it reflects the idea, due to Quine, but also reflected in ordinary practice, that in translation or paraphrase there is considerable latitude. A theory of interpretation is unique only up to the formal and empirical constraints on interpretation, and translation, on this view, is unique only up to a pair of these theories.

If an appropriate equivalence concept can be drawn from empirically applied theories of truth and the theories include Davidson's paratactic semantics for sentences of indirect discourse, then a theory of interpretation

whose central notion is truth will effect the transition from direct to indirect discourse.

The situation is more difficult still with the other attitudes. What do we need to infer (15) from (13), for example? There are obvious objections to extending Davidson's paratactic analysis to expressions like 'believes that'. There may be no utterance of Kurt's to samesay each time we attribute a belief to him. A person may have a belief that he has no occasion to express. This poses a problem for an analysis that rests on speech. An interpreter of a language that includes sentences like (13) must find a general way to bring these sentences under a theory of truth. He needs to see how their semantic character is built up from a finite vocabulary in a finite number of steps. It is worth noting, however, that no general analysis of these sentences is needed to establish the inferences between an attitude towards a sentence and an attitude towards a proposition. The first attitude offers a quoted sentence and so a paratactic treatment of the second attitude is plausible here. When we attribute a belief to a person who accepts a sentence or holds it true, there is a speech act that we can associate with that person and that we can mimic. My utterance of (15) is true if my utterance of 'It is raining' translates some sentence that Kurt holds true.

*5 **Belief and meaning*** A speaker holds a sentence true as a result of two factors: what he takes the sentence to mean and what he believes. In taking truth as the central notion in a theory of interpretation, Davidson assigns two roles to truth: a role in the theory of meaning and a role in the theory of belief. What a person believes is that certain truth conditions are fulfilled. What his sentences mean is that if these conditions are fulfilled, the sentences are true. Kurt holds 'Es regnet' true at t because Kurt knows that an utterance of this sentence at t is true if and only if it is raining near the speaker at t and Kurt believes that it is raining near the speaker at t. The words to the right of the biconditional and the words to the right of the subordinating conjunction are the same words. It is only because the words are the same words that a particular attribution of meaning and belief together offer a reason for Kurt to hold 'Es regnet' true. Davidson remarks that meaning and belief are linked at the level of evidence by the notion of holding a sentence true, but they are linked at the level of theory too; here they are linked by the notion of a truth condition. Truth conditions interpret a person's words and spell out the contents of his attitudes. This resembles an idea of Frege's: what we know when we are able to interpret a sentence is the meaning or sense of the sentence, and this is the object of the belief that explains why we hold the sentence true.

When an interpreter attributes a certain meaning to a speaker's utterances, what he attributes is some knowledge to the speaker: the speaker knows that the utterance is true if and only if certain conditions obtain. He attributes this knowledge to the speaker on the basis of his own observations of the sentences the speaker holds true, the conditions under which he holds them true and on the basis of charity. Charity enters when the interpreter assumes that his beliefs about the conditions that obtain at the time a speaker holds a sentence true are also the speaker's beliefs. The interpreter, as Davidson has said, holds belief constant and solves for meaning.

The acts of holding sentences true that the interpreter will observe are few, but the sentences for which the theory, based on these observations, offers an interpretation are infinite in number. On the basis of the theory, the interpreter will attribute to a speaker knowledge of the truth conditions of sentences that the speaker will never speak or hold true. However, in attributing this knowledge to the speaker, the interpreter implies that if the speaker was, at some time, to hold these sentences true, it would be because of his knowledge of their truth conditions and his belief that these conditions are fulfilled at this time. Davidson's theory is silent on how a speaker acquires these beliefs or how he might justify them. It is this silence that Michael Dummett in his papers "What Is a Theory of Meaning? (I)" (see [5]) and "What Is a Theory of Meaning? (II)" (see [6]) draws our attention to.

At the center of Davidson's approach to meaning and belief is the notion of reason. A theory of understanding or interpretation is a theory in light of which the actions and attitudes of other people can be seen to be reasonable. The idea that we understand or explain a person's actions by rationalizing them has its roots in common practice and in social science. Decision theory and the theory of consumer demand have formalized the idea and drawn it into a theory of independent and dependent decision under perfect and imperfect information, but the idea is still the simple one that a person's action is rational if it is the action that he believed would bring him his fondest desire. Here the agent's beliefs and desires are taken for granted, and it is his action that is shown to be rational or reasonable given these attitudes. The idea of rationality here is thin. No matter how irrational or unreasonable the person's attitudes, his action is rational or reasonable, if, given the attitudes, the action is one that it was reasonable for him to perform. But what of the attitudes themselves? The idea of reasonableness should apply to them as well. After all, given an occasion or a body of prior attitudes, there are some attitudes that it would be reasonable for a person to hold and others that it would not be. A theory of interpretation or understanding that relies on the notion of reason should extend this notion to the attitudes.

The anthropologist A. R. Radcliffe-Brown objected to the rational explanations that Sir James Fraser offered of the ritual actions of the natives of the South Pacific (see [8], pp. 133-152). Fraser explained actions like rain dancing in a quite straightforward manner. The natives, according to Fraser, desired rain and believed that dancing was the best way to get it. On Radcliffe-Brown's view, Fraser was trading one mystery for another. Fraser's explanation of the natives' action attributes to the natives "superstitious" beliefs, beliefs that are radically different from the anthropologist's own. This leaves Fraser with a problem. He has to explain how people of apparently sound mind and body, people who in many respects believe as the anthropologist does, could have come to hold these unreasonable beliefs. Why did the Maori believe that dancing would cause rain when dancing was never or hardly ever followed by rain? Why hadn't years of dancing and drought infirmed the natives' belief? Fraser's theory is silent on these questions. It does not explain the difference between the attitudes of the native and the anthropologist, and it does not explain how the natives' unreasonable beliefs cohere with their more reasonable ones or how these unreasonable beliefs could have been acquired.

The objections that Dummett raises to Davidson's program are like the objections that Radcliffe-Brown raised against Fraser's. A theory of interpretation must not only rationalize action, it must rationalize belief as well. A theory of rational belief for a person should show how his beliefs are based on his experience. The theory should start with his accumulated observations and project them by means of a theory of confirmation onto his total network of belief. Fraser offered no such theory. Dummett doesn't believe that Davidson has either. As he sees it, Davidson, in taking truth as the central notion, tailors his theory of interpretation to meaning, and, as a result, does not and cannot offer a full account of reasonable belief.

On what basis, for example, would it be reasonable for a speaker to believe that the truth conditions of an occasion sentence are fulfilled? One answer that an interpreter could give is that observation provides a basis. On the basis of his own observations, the interpreter believes that a certain condition, *C*, obtains whenever the speaker holds the sentence 'Es regnet' true. Charity invites the interpreter to attribute to the speaker the belief that *C* obtains on each occasion that he holds this sentence true. In this case, the interpreter has a procedure for determining whether the truth conditions of the sentences are fulfilled, viz., observation, and charity allows him to attribute this procedure to the speaker as well.

However, even occasion sentences that the interpreter observes speakers holding true pose problems of reasonable belief. Because of the exceptions to any generalization about the conditions under which speakers hold a sentence true, on some occasions when a speaker holds a sentence true, the interpreter will want to say that the speaker has made a mistake. A speaker makes a mistake when he holds a sentence true under circumstances that do not fulfill the truth conditions associated with the sentence; the speaker believes that the conditions are fulfilled, but the interpreter believes that they are not. Once Davidson's interpreter decides that a speaker or most speakers associate a particular condition, *C*, with a particular sentence, *s*, he is able to suspend the principle of charity and able to speak of mistakes. He decides this on holistic grounds: the totality of *T*-sentences that optimally fit the evidence about sentences held true includes a *T*-sentence that associates *C* and *s*. Here the interpreter holds meaning constant and solves for belief.

The difficulty is that now the interpreter assigns a belief to the speaker that differs radically from his own, and Davidson's theory leaves this difference in belief entirely unaccounted for. As Dummett says, the theory offers no explanation of how such mistakes occur; it offers us no explanation of how the speaker acquired this 'false belief' or how he might proceed to justify the belief (see [5], p. 120). Dummett argues that a theory which fails to offer this has no right to appeal to the notion of a mistake and that a theory of meaning which takes truth rather than verification as its central notion leaves no room for mistakes.

This consideration of mistakes leads to a larger point. When there is no reason to cite a mistake, when the interpreter can explain the fact that the speaker holds a sentence true on the assumption that the speaker believes as he does, the origin or basis for the speaker's belief is still unaccounted for. The speaker is credited with the belief because of the principle of charity,

but to say this is not enough. A theory of interpretation should reveal the grounds the speaker has for adopting such a belief. When his belief is our belief, we overlook this; we assume that his grounds are our grounds. When the belief we attribute to him is different from ours, this assumption is not possible, and the need to say something about the origin or basis of his belief and our evidence for this becomes apparent. In short, Davidson starts with the fact that a person holds a sentence true and explains how this depends on his knowing truth conditions and believing that these conditions are fulfilled, but he does not explain how a person might acquire his beliefs or his knowledge.

The justification becomes more acute when we turn to standing sentences. These sentences are not keyed to the occasions when speakers hold them true and, consequently, there is no belief of the interpreter's about these occasions that he can charitably and usefully attribute to the speakers when they do hold them true. Nevertheless, the interpreter may have beliefs that he can attribute to the speaker that will explain why the speaker holds a standing sentence true. Suppose, for example, that Kurt holds 'Der Schnee ist weiss' true. The interpreter explains this by attributing to Kurt the knowledge that 'Der Schnee ist weiss' is true in German if and only if snow is white and the belief that snow is white. On what basis does the interpreter attribute these to Kurt? On the basis of the pattern of sentences held true and charity. Here, however, there is less to charity. The interpreter's own grounds for believing that snow is white are somewhat less obvious, and, consequently, he has less reason to believe that his grounds for believing that snow is white are Kurt's grounds.

According to Dummett, the sentences that raise the most serious difficulty are the undecidable sentences. Dummett believes that subjunctive conditionals and sentences that quantify over infinite domains are not in principle decidable. The ability to recognize these sentences as true or to recognize them as false extends beyond any person's capacity. There is no process whereby a person might come to accept one of these sentences as having been conclusively established as true: no process of observation or inference. Given a theory of interpretation with truth as the central notion, the meaning of an undecidable sentence, like any other, is its truth condition. A person who understands or knows the meaning of the sentence knows the conditions under which the sentence is true. However, if the sentence is undecidable, neither the speaker nor the interpreter can ever have any reason for believing that these conditions are fulfilled. According to Dummett, this removes any reason an interpreter might have for attributing knowledge of these truth conditions to the speaker. There is no belief here that speaker and interpreter share and no basis on which to hold belief constant and solve for truth conditions. As a result, there is no basis on which to identify the ability to understand or interpret these sentences with knowledge of truth conditions.

Dummett's argument seems to boil down to this. It is appropriate to attribute knowledge of a truth condition to a person only if the belief that the condition is fulfilled is a belief that the person could acquire through a reasonable procedure of observation or inference. This is Dummett's objection

to "realism". Realism is the doctrine that the truth conditions of a sentence may obtain or not irrespective of any person's capacity to know that they do. Dummett objects that we cannot credit speakers with a conception of truth that is independent of their capacity to know whether the conditions for truth are fulfilled. We cannot credit them with such a conception because there is no way they could have acquired it. The conclusion that Dummett draws from this is that we do not possess a concept of truth applicable to all sentences, and hence knowledge of the truth conditions of a sentence cannot be a condition for understanding it.

What Dummett's objections draw our attention to is how much better truth serves meaning than it serves belief. Charity fixes belief, and once belief is fixed, truth conditions offer a theory of meaning. But if we had taken meaning for granted, could we have used truth conditions for a theory of reasonable belief? The idea would be that every belief is a belief that a certain truth condition is fulfilled, and every time a person holds a sentence true, he believes that the truth conditions of that sentence are fulfilled.

The difficulty is that, on this account, there are no principles or constraints that assure that any of these beliefs fall within reason. A person's beliefs, like his language, form an articulated structure, with some beliefs lying at the periphery and others lying further away and within the interior. Beliefs at the periphery are reasonable in light of a person's direct experience or observation, beliefs lying within the interior are reasonable in light of an acceptable line of inference from other beliefs that are themselves reasonably held. Reasonable belief, in the first instance, is a property only of beliefs at the periphery. However, because beliefs that form the structure are interconnected, reasonableness is transmitted from the periphery to the interior.

A comprehensive theory of meaning and belief must elaborate the structure of both meaning and belief. Given the terms of the theory, the beliefs that an interpreter attributes to a speaker must be beliefs that lie at the periphery or that are connected in an appropriate way to those that do. The trouble with truth conditions is that, in many cases, no line of inference can be drawn from the periphery to the belief that the truth conditions are fulfilled. This weakens the claim that a theory of meaning with truth as its central notion leads to a comprehensive theory of human understanding. A theory that leaves a person's beliefs unexplained provides, at best, an incomplete understanding of his thoughts and his actions.

REFERENCES

[1] Davidson, D., "On saying that," *Synthese*, vol. 19 (1968-1969), pp. 130-146.

[2] Davidson, D., "Radical interpretation," *Dialectica*, vol. 27 (1973), pp. 313-328.

[3] Davidson, D., "Belief and the basis of meaning," *Synthese*, vol. 27 (1974), pp. 309-323.

[4] Davidson, D., "Action and reaction," *Inquiry*, vol. 13 (1980), pp. 140-148.

[5] Dummett, M., "What is a theory of meaning? (I)," in *Mind and Language*, ed., S. Guttenplan, Oxford University Press, Oxford, England, 1975.

[6] Dummett, M., "What is a theory of meaning? (II)," pp. 67-137 in *Truth and Meaning*, eds., G. Evans and J. McDowell, Oxford University Press, Oxford, England, 1976.

[7] Dummett, M., "The significance of Quine's indeterminacy thesis," *Synthese*, vol. 27 (1974), pp. 351-397 (reprinted in *Truth and Other Enigmas*, Harvard University Press, Cambridge, Massachusetts, 1978).

[8] Radcliffe-Brown, A. R., *Structure and Function in Primitive Society*, Free Press, New York, New York, 1965.

Department of Philosophy
University of Minnesota
Minneapolis, Minnesota 55455