

## **DUAL ASPECT SEMANTICS**

Frege's notion of sense plays (at least) two roles in his theory of meaning. One role concerns the relation between language and reality: an expression's sense determines its reference. The other role relates a language to the mind of someone who understands it: to understand an expression is to grasp its sense. The dual role of sense is seen clearly in Frege's account of the semantics of identity statements. 'The morning star = the evening star' is true, since the sense of 'the morning star' and the sense of 'the evening star' determine the same reference. The sentence is cognitively significant, since it is possible for someone to know the senses expressed by the expressions 'the morning star' and 'the evening star' yet not know that they determine the same reference.

During the last fifteen years or so there has been a sustained attack on the Fregean conception of sense. An examination of proper names, indexicals, and natural kind terms, has led many philosophers of language and mind to conclude that no single notion of sense can play both roles. Hilary Putnam puts the point by saying that "no theory can make it the case that 'meanings' are in the head and simultaneously make fit the case that 'meanings' determine external world reference" (Putnam 1986, p. 12). Of course for Frege, meanings, i.e., senses, are in the head (in that they are grasped) and determine reference.

A number of philosophers have responded to these arguments by constructing two-tiered or dual-aspect theories of meaning. We will call them DATs (Block, 1985; Field, 1977; Fodor, 1980; Harman, 1973, 1974, 1982; Loar, 1982; Lycan, 1981, 1982a, 1982b, 1984; McGinn, 1982). According to these accounts, a theory of meaning for a language L consists of two distinct components. One component is intended to provide an account of the relations between language and the world: truth, reference, satisfaction, etc. The other is supposed to provide an account of understanding and cognitive significance. In this paper we will examine a particular proposal concerning the appropriate form of a DAT according to which the two components are:

I. A theory of reference and truth for L, and

II. A characterization of the conceptual roles of sentences and other expressions of L.

We will contrast DATs with an approach which is like Frege's in one important respect: it employs a single notion to serve both the purpose of the theory of reference and the theory of understanding. Its central tenet is that a theory of meaning for L is a certain kind of truth theory for L. Since Donald Davidson is the most prominent and subtle defender of this approach we will call such theories of meaning 'Davidsonian truth theories' (Davidson, 1967, 1973b, 1974;

Lepore, 1982b, 1983; Loewer, 1982; Lepore and Loewer, 1981, 1983). At first it may seem that this approach is contained in a DAT, since the latter has a truth theory as one of its components. But this is not so. We will show that the truth theory component of a DAT is quite different from that of a Davidsonian truth theory. We will argue that, by separating a theory of meaning into a theory of reference and a theory of conceptual role, DATs are unable to serve as theories of interpretation or as accounts of cognitive significance.

The organization of our paper is this: we first examine the problems that motivate the construction of DATs. Then we discuss the form of a DAT, focusing primarily on a proposal due to Colin McGinn (1982). We then develop a Davidsonian theory of meaning showing how truth theories serve as theories of interpretation. As such, they provide both an account of truth and reference, on the one hand, and an account of understanding and cognitive significance, on the other. In the next section we show that DATs do not make adequate theories of interpretation. We also challenge their adequacy as theories of meaning for languages of thought. Finally, we return to the problems motivating DATs and discuss the extent to which they can be accommodated within a Davidsonian framework.

### **MOTIVATION FOR DAT**

Hilary Putnam asks us to imagine two planets, Earth and Twin-Earth, and two of their residents, say, Arabella and twin-Arabella. Twin-Earth is almost a physical replica of Earth. The only

difference is that on Twin-Earth the clear liquid the twin-people drink, that fills their oceans, and that they call 'water', is composed not of H<sub>2</sub>O molecules but of XYZ molecules (Putnam, 1975).

According to Putnam, the expression 'water' on Earth refers to the stuff composed of H<sub>2</sub>O and not composed of XYZ. It is exactly the reverse for the expression 'water' on Twin-Earth. This is so even if no speakers of English and twin-English know the molecular structures of water and twin-water or can distinguish between the two. Putnam argues as follows: in Frege's theory to understand an expression, say, 'water', is to 'grasp' its sense. Exactly what it is to grasp a sense is not all that clear, but it is to be in some psychological state or other; perhaps the state of believing that 'water' expresses a certain sense. Since Arabella and twin-Arabella are physically type identical,<sup>1</sup> they are in type identical psychological states. So if each understands her word 'water', then each grasps the same sense. But the references of their words differ. Putnam concludes that if sense is what is grasped when understanding an expression, then sense does not determine reference. If sense is what determines reference, then sense is not what is grasped in understanding.

Putnam's initial reaction to this argument was to distinguish two components of meaning. One he calls 'stereotype'. It is the information which linguistically competent speakers associate with an expression. The stereotype of 'water' as used both on Earth and on

Twin-Earth consists in the information that water is a clear liquid, that quenches thirst, that fills the oceans, and so on. To understand 'water' is to know its grammatical role and its stereotype. This is supposed to be the 'mind' component of meaning (Putnam, 1975).

The second component of meaning is reference. On Putnam's account the reference of a natural kind expression like 'water' is determined by facts which are outside the minds of users of the expression. For example, 'water' refers to whatever stuff is structurally similar to this stuff (pointing at samples of water). Given that water is H<sub>2</sub>O (and that the relevant kind of structural similarity is sameness of chemical composition), the extension of 'water' on Earth is H<sub>2</sub>O. Analogously, the extension of 'water' on Twin-Earth is XYZ. On Putnam's theory the stereotype of an expression is the mind component of sense, its reference is the world component and, as the Twin-Earth story shows, the first does not determine the second. The theory of meaning thus divides into two parts: a theory of understanding (and cognitive significance) and a theory of reference (and truth).

Putnam's argument for the bifurcation of meaning depends on accepting his view that 'water' on Earth refers to H<sub>2</sub>O, while 'water' on Twin-Earth refers to XYZ. This is a claim which can be (and has been) disputed (Zemach, 1976). But even if Putnam is mistaken about the semantics of natural kind terms, there are other examples that lead to dual component views.

Imagine that Arabella and twin-Arabella each utters 'I am 30 years old'. Once again they are in identical psychological states, but the references of their utterances of 'I' differ, and even the truth values of their utterances can differ. They are physical replicas, but Arabella came into existence only a few minutes ago. Arabella's and twin-Arabella's understandings of 'I' are the same, although their references and the truth values of their utterances differ. This shows that if the sense of 'I' is what is grasped by a person who understands it then that sense does not by itself determine reference. On the other hand, if sense determines reference then Arabella and twin-Arabella do not grasp the same sense. David Kaplan, among others (Kaplan, 1977; cf also Perry, 1978, 1979; White, 1983), distinguishes the character of an expression from its content in a context. The character of an expression is a function from contexts of utterances to contents, e.g., the character of 'I', according to Kaplan, maps a context of utterance onto the utterer. It is the character of 'I am 30 years old' that is grasped by someone who understands the sentence. The utterance's content is its truth conditions. When Arabella and twin-Arabella each utter 'I am 30 years old', what is in their minds may be the same (they have the same understanding of the sentence) but their utterances have different contents and so may differ in truth value. Kaplan's account, like Putnam's, is a two tiered theory of meaning. But it differs from Putnam's in that sense is unlike character in that it does not determine content relative to context, at least as 'context' of utterance is normally construed (see White, 1983).

Our discussion so far seems to show that two expressions can have the same stereotype, or character (or, whatever corresponds to cognitive significance), and yet possess tokens which differ with respect to reference and truth conditions. It has been argued, conversely, that sentences with the same truth conditions can differ with respect to the understanding component of meaning (Kripke, 1979). According to Kripke proper names designate rigidly (Kripke, 1972). It has been claimed that it follows from this that, for example, the truth conditions of the sentences 'Cicero is bald' and 'Tully is bald' are identical. But what are we to make of Arabella, who understands both sentences and assents to the first and dissents from the second? If understanding a sentence is knowing its truth conditions, then it follows that Arabella is flatly contradicting herself, since she is asserting and denying statements with identical truth conditions. Furthermore, it seems to follow that she has contradictory beliefs: she believes that Cicero is bald and believes that Cicero is not bald. But these are not ordinary contradictions, since no amount of thought on her part would enable her to recognize that she has contradictory beliefs. William Lycan (1982; see also Lycan, 1985, pp. 90-91) reacts to this problem by saying:

Nothing that [Arabella] carries in her head enables her to tell that ... 'Cicero' and 'Tully' represent the same person. And, therefore, there is no way for her to deduce from her mental machinery anything she could recognize as a contradiction. The names 'Cicero' and 'Tully' obviously play

distinct computational roles for [Arabella] ....

Lycan intends this as a solution to the problem as it arises for mental representations. His idea is that there are two distinct ways of semantically individuating Arabella's mental representations; according to truth conditions and according to computational role. According to the way of truth conditions, Arabella believes that Cicero is bald and also believes that Cicero is not bald. But this does not impugn her rationality, since truth conditions are not part of Arabella's 'mental machinery'. According to the way of computational role, she believes that Cicero is bald and that Tully is bald are distinct, since her representations 'Cicero is bald' and 'Tully is bald' have different computational roles. When her beliefs are individuated in terms of computational role, Arabella does not have contradictory beliefs.

We could extend Lycan's account (though he does not make this extension) to the semantics of natural languages if we could find something to play the part of computational role for natural language expressions. The simplest suggestion is that the computational role of a person *P*'s sentence *S* at time *t* is the same as the computational role of the mental representation constituent of the belief expressed by *S* for *P* at *t*. 'Cicero is bald' and 'Tully is bald' have the same truth conditions in English, but they may differ in their computational roles for a particular speaker at *t*. It is computational role that characterizes one's understanding of an

expression. We can see how a person might understand both 'Cicero is bald' and 'Tully is bald', and assert one and deny the other, even though they have the same truth conditions.

As our discussion of Lycan's proposal makes clear, DATs have been proposed for mental representations, a.k.a. languages of thought, as well as for natural languages. Jerry Fodor, the principal proponent of languages of thought, has been developing a theory of mental states and processes he calls the computational theory of mind (CTM). According to CTM, mental states and processes are computations over representations. For example, believing that snow is white is being in a certain computational relation to a representation which means that snow is white. The system of mental representations is like a language in that representations possess both a syntax and a semantics. It is a central tenet of CTM that computations apply to representations in virtue of their syntactic features (Fodor, 1980, p. 226). While Fodor admits that it is not all that clear what count as syntactic features he is clear that semantic properties, e.g., truth and reference, are not syntactic. The mind (and its components) has no way of recognizing the reference or truth conditions of the representations it operates on. Instead, it operates on syntactic features of representations which 'represent' the semantic features. The computational role of a mental representation must depend upon, and only upon, those properties of representations which do not advert to matters outside the agent's head (McGinn, 1982, p. 208).

Fodor claims that a consequence of CTM is a formality condition, which specifies that in CTM psychological states count as different states only if they differ computationally. Applied to belief, this means that beliefs can differ in content only if they contain formally distinct representations. This supervenience principle, that *S* and *S'* are distinct psychological states only if they are distinct computationally, lies at the heart of CTM. Although Fodor endorses the formality condition, he also thinks that cognitive psychology contains true generalizations connecting propositional attitudes with each other, environmental conditions, and behavior. An example of the sort of generalization he has in mind, is: if someone wants to go downtown and believes that the bus provides the only way to get there, then, *ceteris paribus*, he will take the bus. As Fodor emphasizes, the specification of propositional contents in these generalizations is essential to their explanatory role. It is a person's belief *that the bus provides the only way to get downtown that* explains his taking the bus. At first, this may seem incompatible with the claim that only formal properties of representations are relevant to the computations which produce behavior. However, there is no incompatibility as long as the contents of attitudes are specified in a way that respects the formality condition. This means that two representations can differ in content only if they differ syntactically. Fodor observes that a characterization of meaning which conforms to the formality condition is methodologically solipsistic (Putnam, 1975) in that differences of meaning depend entirely

upon internal mental characteristics, e.g., computations over representations.

We have described Fodor's views at some length because we want to show why a DAT theory seems to fit the bill as a theory of meaning for languages of thought. Fodor observes that truth conditional semantics for a language of thought is not methodologically solipsistic. It fails to conform to the formality condition (Fodor, 1982, p. 22). Putnam's Twin-Earth examples show this. Arabella and twin-Arabella are computationally identical when each is thinking what each would express by uttering 'Water is wet'. So, each bears the same computational relation to formally identical representations. But the truth condition of the token representation in Arabella's mind is that H<sub>2</sub>O is wet, while the truth condition of the representation in twin-Arabella's mind is that XYZ is wet. There is a difference in truth conditions without a corresponding difference in formal properties. The characterization of contents in terms of truth conditions may seem defective from the perspective of CTM in another way as well. 'Water is wet' and 'H<sub>2</sub>O is wet' are claimed to have the same truth conditions, but certainly there is a difference between believing that water is wet and believing that H<sub>2</sub>O is wet. Truth conditions seem to be both too fine grained (the Twin-Earth problems) and too coarse grained (Kripke's puzzle (Kripke, 1979)) to specify the contents of mental representations.

It should be clear why DATs have been proposed as theories of meaning for languages of thought. The truth conditional component of a DAT characterizes the relations between

representations and the world. But a second component is needed which characterizes content in a way that conforms to the formality condition and is fine grained enough to capture differences in belief like the one mentioned a paragraph back. This second component is the mind-component aspect.

#### THE FORM OF DAT

So far, we have discussed some issues which motivate a distinction between two aspects of meaning. Some philosophers have claimed that the correct way for a theory of meaning to accommodate the two aspects is by containing two autonomous components, a truth conditional component, and a component accounting for the use or understanding features of meaning.

Colin McGinn (1982, p. 229) explicitly advocates such a view.

For perspicuity we can separate out the two contributions by taking the meaning ascription as equivalent to a conjunction: For  $S$  to mean that  $p$  is for  $S$  to be true iff  $Q$  for some ' $Q$ ' having the same truth conditions as ' $p$ ', and for  $S$  to have some cognitive role  $\phi$  such that ' $p$ ' also has cognitive role  $\phi$ . ... Now to have a complete theory of meaning would be to have adequate theories corresponding to each conjunct of this schema.

McGinn is not only claiming that an adequate theory of meaning consists of two separate

theories, but is also offering an analysis of '*S* means that *p*'.

The first component, the truth theory, may seem relatively unproblematic. It is supposed to entail, for each sentence *S* of language *L*, an instance of:

(T) *S* is true in *L* iff *p*,

where '*S*' is replaced by a structural description of a sentence of *L* and '*p*' is replaced by a meta-language sentence which specifies *S*'s truth conditions. Tarski required that the sentence replacing '*p*' be a translation of the sentence replacing '*S*' (Tarski, 1956). Putnam, Field, Fodor, and McGinn do not have this conception in mind. Fodor says (Fodor, unpublished) that "a truth condition is an actual or possible state of affairs. If *S* is the truth condition of (the formula) *F*, then *F* is true iff *S* is actual." According to McGinn, "...a truth theory is a specification of the facts stated by sentences of the object language, in the intuitive sense of that recalcitrant notion" (McGinn, 1982, p. 232). As McGinn says, 'fact' is a recalcitrant notion. 'State of affairs' is no clearer. However, it is clear that some advocates of DATs would count 'Water is wet' (uttered by an English speaker) as stating the same fact or state of affairs as 'H<sub>2</sub>O is wet', and 'Tully is bald' as stating the same fact or state of affairs as 'Cicero is bald'. So, their characterization of an adequate truth theory is different from Tarski's, since two sentences can state the same fact (or

state of affairs) without being good translations of each other. Of course, this is not a worry to advocates of DATs, since sameness of meaning requires not only sameness of truth conditions but sameness of cognitive role as well.

The second component of McGinn's DAT is a theory of cognitive role. Other authors use the terms 'conceptual role' (which we prefer) and 'computational role' for similar, though perhaps not identical, notions. While the idea of conceptual role has been around for a while, the form of a theory of conceptual role is much less clear than the form of a truth theory. Sellars speaks of two sentences having the same conceptual role if they are related by inference, both deductive and inductive, to the same sentences in the same ways. Sellars also includes relations between sentences and perception and action, 'language entry and exit rules', in his specification of conceptual role (Sellars, 1956, 1963, 1969). Harman, thinking of the language of thought, characterizes the conceptual role of an expression by its relations to perception, to other expressions, and to behavior (Harman, 1973, 1974, 1982). Both Sellars' and Harman's characterizations suggest that conceptual role theories for a language  $L$  will take the form of a theory of inference for  $L$ , combined with a causal theory of perceptual inputs and outputs. But neither provide detailed accounts of these theories.

McGinn relies on Hartry Field's account of conceptual role (Field, 1977). Field characterizes conceptual role in terms of a probability function defined over all the sentences of

a person's language. It specifies a person's commitments concerning how he will change his degrees of belief when he acquires new information. The probability function, by specifying inductive and deductive relations, characterizes the conceptual roles of expressions. A and B are said to have the same conceptual role iff  $P(A/C) = P(B/C)$  for all sentences C in the language. On this account 'Tully is bald' and 'Cicero is bald' may have different conceptual roles for a person, since there may be an S for which  $P('Tully is bald'/S) \neq P('Cicero is bald'/S)$ . The conceptual role of a non-sentential expression is specified in terms of the conceptual roles of all the sentences in which it appears. There may be simple characterizations of the conceptual roles of some expressions. For example, the role of negation is specified by the probability laws involving negation.

McGinn claims that two component theories, containing a truth theory and a Fieldian conceptual role theory, can deal with the problems we discussed in the first section (McGinn, 1982, pp. 234-237, 247). Consider Arabella and twin-Arabella. Their languages, English and twin-English, are syntactically identical, and since the twins are physically type identical, the conceptual roles of their expressions are isomorphic. Each one's sentence 'Water is wet' has the same (or isomorphic) conceptual role(s) so their mental states are identical. However, their sentences differ in their truth conditions since the fact that makes Arabella's sentence true is H<sub>2</sub>O's being wet, while the fact that makes twin-Arabella's sentence true is XYZ's being wet. A

similar remark can be made concerning indexicals. The sentence 'I am 30 years old' has the same conceptual role for Arabella and twin-Arabella, but tokens of the two differ in their truth conditions. Since it is the conceptual role of an expression (in the language of thought) which determines its role in the production of behavior, Arabella and twin-Arabella will behave identically when each believes what she would express by saying 'I am 30 years old'. But since conceptual role does not determine truth conditions, the truth values of their beliefs may differ.<sup>2</sup> The dual component view also seems able to account for sentences which apparently have the same truth conditions but differ in meaning. Thus, 'Cicero is bald' and 'Tully is bald' are supposed to have the same truth condition but a given speaker's probability assignment might contain an  $S$  such that  $P(\text{'Tully is bald'}/S) \neq P(\text{'Cicero is bald'}/S)$ .

There are a number of features of the Field-McGinn characterization of conceptual role that are worth noting.

1. McGinn's account differs from the kinds of accounts suggested by Sellars, Harman, and Block in which conceptual role is characterized in terms of the *causal* relations that hold among representational mental states, perceptions, and behaviors. An individual's probability assignment does not specify causal relations, but rather his commitments concerning rational change of belief. Only if the probability assignment is reflected in causal relations among belief states, etc., will conceptual role be capable of functioning in psychological explanations.

2. Field's (1978) account of conceptual role obviously involves a great deal of idealization. Gilbert Harman has argued that it is unrealistic to suppose that an individual reasons in terms of probabilities, since this would require keeping track of an enormous amount of information and require an enormous number of computations. Perhaps this objection can be met (Jeffrey, 1983), but a more difficult problem is presented by the evidence which shows that our beliefs do not conform to probability theory. For example, people will often assign a higher probability to a conjunction than to either of its conjuncts and do not typically change beliefs in accordance with conditionalization. So, McGinn's theory of conceptual role might not apply to human thought.

3. Conceptual role is a holistic notion. In characterizing the conceptual role of a sentence one must simultaneously characterize the conceptual roles of all other sentences. Any change in the probability function - even just extending it to a new vocabulary - results in a change in conceptual role for every sentence. Because of this two people will seldom assign the same conceptual role to syntactically identical expressions. Field explicitly offers conceptual role only as an account of intro-individual meaning. He does not think that it makes sense to compare different individual's conceptual roles (Field, 1977). However, McGinn apparently does think that it is meaningful to make interpersonal comparisons of conceptual role. For example, he speaks of Arabella and her twin's mental representations as having the same

conceptual roles.

4. Field's characterization of conceptual role is solipsistic, since it is characterized entirely in terms of ingredients within the mind of the individual. It is this feature which suggests to McGinn that conceptual role can provide an account of the aspects of meaning that meets Fodor's methodological solipsism constraint. On some other versions, e.g., Harman's, the characterization of conceptual role also includes relations among sentences, environmental features, and behavior. So, Harman's account of conceptual role, as he insists, is non-solipsistic (Harman, 1986). We can imagine a theory between Field's and Harman's (along this dimension) which includes relations to sensory inputs and behavioral outputs in the characterization of conceptual role. As long as the inputs and outputs are described in ways that do not entail the existence of anything other than the thinker's body, the solipsistic nature of the account is preserved. The difference will be important when we come to evaluate the adequacy of conceptual role theories as semantic theories for languages of thought.

5. It should be clear that Putnam's stereotype and Kaplan's character are quite different from conceptual role (and from each other). The stereotype of 'water' is the information which a typical competent speaker of English associates with 'water', e.g., the water is liquid, necessary for life, fills the oceans, etc. Stereotype differs from conceptual role in a number of ways. (a) Stereotype characterizes cognitive significance, since it specifies the information associated

with a term. It is not obvious that the conceptual role of a term or sentence associates with it any information (see fourth section). (b) It is not clear that stereotype is 'in the head' in the way conceptual role is. The expressions used to characterize the information contained in stereotype are themselves subject to Twin-Earth arguments, and this seems to show that stereotype itself is not solipsistic. (c) As we have pointed out, conceptual role is holistic. Stereotype does not appear to be holistic.

There are also important differences between conceptual role and character. (a) Two people can associate the same character with 'I am hungry', even though the sentence has different conceptual roles for each, since the two may differ in their overall probability assignments. So, character can be used to explain the sense in which two people who assert 'I am hungry' share the same belief, while conceptual role cannot. (b) Character determines truth conditions relative to context, but there is no systematic relation between conceptual role, context, and truth conditions. At least, none is built into McGinn's account.

However McGinn might fill in the details of his dual component view, the general picture is clear. On his view the appropriate form for a semantic theory for a language is a conjunction of *two* theories. One characterizing internal mental features of meaning and the other characterizing relations between language and the world. In opposition to the dual component view are semantic theories which provide a *unified* treatment of the mind and world

aspects of meaning. Frege's and Davidson's theories are examples. McGinn, of course, thinks that such unified accounts are misguided. He says, "But it seems that nothing of critical importance would be lost, and some philosophical clarity gained, if we were to replace in our theory of meaning, the ordinary undifferentiated notion of content by the separate and distinct components exhibited by the conjunctive paraphrase" (McGinn, 1982, p. 229). We will argue in the fourth section that, contrary to McGinn's claim, something of critical importance is missed by bifurcating the theory of meaning in the way McGinn proposes. Dual component theories cannot be used as theories of interpretation, and for this reason fail to provide adequate accounts of communication. We will present this argument in the fourth section. First, we want to show how Davidsonian truth theories can be used as theories of interpretation.

#### TRUTH CONDITIONAL THEORIES OF COMMUNICATION AND UNDERSTANDING

According to DATs, a theory of truth for a language is incomplete qua theory of meaning, because it fails to provide an account of the mental aspects of meaning: language understanding and cognitive significance. The conceptual role component is supposed to do that job. This view of the place of a truth theory in an account of meaning is clearly at variance with Donald Davidson's. Davidson sees a truth theory as capable of providing both an account of language understanding and an account of the relations between language and reality. In this section we will show how it is that knowledge of Davidsonian truth conditions can play a central role in

understanding and communication. Our argument is a bit different from the arguments found in Davidson, but we clearly take our cue from his writings.

It is almost a truism (or was a truism until recently) in philosophy of language that to understand a sentence is to know its truth conditions. But if it is a truism, it is an obscure one. We will try to show what truth it contains. Once again, consider Arabella: she utters the words 'Es schneit' within earshot of Barbarella and Cinderella.<sup>3</sup> Barbarella understands German while Cinderella does not. This makes a difference. Barbarella acquires the beliefs that it is snowing and that Arabella believes that it is snowing and perhaps some other beliefs as well. Cinderella does not acquire these beliefs; Arabella's utterances are so many sounds to her. Even if she recognizes them as an assertion, something she is able to do without understanding German, she may acquire only the belief that Arabella's utterance 'Es schneit', whatever it may mean, is true, and perhaps also the belief that Arabella holds her utterance to be true. Still, she does not know what Arabella expresses or believes.

We have argued (Lepore, 1982b; Loewer, 1982; Lepore and Loewer, 1982) that *a theory of meaning for a language L should include information such that someone who possesses this information is, given his other cognitive capacities, able to understand L.* Understanding a language involves many complex abilities, e.g., to respond appropriately to assertions, orders, questions, and so forth. We will focus on one central ability, the ability to

justifiably acquire beliefs about the world, and about what a speaker believes in the presence of that speaker's assertions. Since Cinderella, who does not understand German, can come to know that Arabella's utterance 'Es schneit' is true, we can ask what additional information could enable her to justifiably acquire the beliefs which Barbarella acquires?

A plausible (indeed, we think the inevitable) answer to our question is that if Cinderella knew that 'Es schneit' is true (in German) iff it is snowing, she would be in a position to acquire the target beliefs. Reasoning justifying these beliefs could go as follows:

Paradigm (I)

1. Arabella's utterance 'Es schneit' is true.
2. 'Es schneit' is true iff it is snowing.

So,

3. It is snowing.

Paradigm (II)

4. Arabella believes 'Es schneit' is true.
5. Arabella believes 'Es schneit' is true iff it is snowing.

So,

6. Arabella believes that it is snowing.

Elsewhere we have argued that such reasoning gives substance to the claim that to understand a sentence is to know its truth conditions (Lepore, 1982b, 1983; Lepore and Loewer, 1981, 1983). However, this claim requires some qualification. We are not saying that a person's understanding of German involves his going through the above inferences, or even that every person who understands German explicitly knows the truth conditions of German sentences. Our claim is that truth conditions explicitly state information which can be used (usually together with other information) to interpret utterances. In this way, a specification of truth conditions for a language can provide an illuminating characterization of language understanding and communication.

The view that a theory of truth for  $L$  can serve as a theory of meaning for  $L$  is most prominently associated with Donald Davidson. Our two paradigms exhibit exactly how a theory of truth can play the role of a theory of meaning. But, of course, not just any theory which entails for each indicative sentence  $S$  of  $L$  a theorem of the form ' $S$  is true iff  $p$ ' can serve as a theory of interpretation for  $L$ . According to Davidson, a theory of interpretation for an individual's language should assign truth conditions to his utterances in a way that results in an attribution of beliefs and preferences to him which are reasonable given his situation and

behavior. Exactly what we count as reasonable will depend on our theories of belief and desire acquisition and our theories of behaviour, etc. (Davidson, 1973b, 1974). We would add that a theory of truth cum theory of interpretation should yield theorems which can be employed in our paradigms. While we have not developed the adequacy conditions on Davidsonian truth theories in detail, what we have said is sufficient to distinguish among truth theories for L. For example, they may serve to eliminate theories which entail "'Schnee ist weiss" is true iff grass is green'. If we use a theory which contains this theorem then we might infer from Arabella uttering 'Schnee ist weiss' that grass is green and that she believes that grass is green. But Arabella might not have this belief. In any case, we would not be justified in believing that she has this belief (or that grass is green) on the basis of the truth of her utterance. It is interesting to see that our constraints also distinguish the truth conditions "'Water is wet" is true iff water is wet' from "'Water is wet" iff H<sub>2</sub>O is wet'. The latter, but not the former, license an inference from Arabella who believes that her utterance 'Water is wet' is true to Arabella believes that H<sub>2</sub>O is wet. We can imagine circumstances in which this would lead to error, i.e., when Arabella fails to believe that water is H<sub>2</sub>O.

Readers familiar with discussions of Davidson's accounts of language will notice that we have emphasized the importance of knowledge of truth conditions, while saying little concerning the nature of the theory which implies instances of (T). Some writers, for example,

Harman, claim that whatever a truth theory has to say about meaning is contained in the recursion clauses of the theory which show how truth conditions of complex sentences depend on semantic features of their component expressions (Harman, 1974; cf. also Fodor, 1975). Harman argues that such a theory at best characterizes the meanings of logical constants, 'and', 'or' etc., by characterizing their conceptual roles, but that the theory does nothing to specify the meanings of other expressions. It should be clear that we disagree (cf. Lepore and Loewer, 1981). Truth conditions do specify meanings in that they enable someone who knows the truth conditions of sentences to interpret the speech of another. Of course, the theory is important as well, but not because it characterizes the conceptual roles of the logical connectives. (It is not clear that it does. Cf. Lepore, 1982a.) Having a truth theory for *L* is important because it provides a specification of truth conditions for all the (infinitely many) sentences of *L* in a way that does not presuppose an understanding of *L*.

We claim that truth theories for natural languages which are theories of interpretation address both aspects which concern DATs. This is clearly seen from our paradigm inference patterns. Someone who knows the truth conditions of the sentences of a language, and knows that this is common knowledge among speakers of the language, is in a position to draw conclusions about the world and about what other speakers have in mind. On the one hand, truth conditions relate sentences to the world. They specify what must hold for a sentence to be true.

On the other hand, they specify what is known by someone who understands a language. Whether they can deal with the specific problems that motivated DATs remains to be seen. We will discuss this matter in the last section.

#### WHAT'S WRONG WITH DAT

In this section we will compare DATs, focusing mainly on McGinn's version, with Davidsonian truth theories. We will argue for three claims: (1) McGinn's account of '*S* means that *p*' involves necessary conditions for sameness of meaning which are much too restrictive and render it incapable of providing an account of communication. (2) Neither component of a DAT, nor the two together, is a theory of interpretation. (3) The conceptual role component of a DAT does not supply the sort of semantics for the language of thought that is required by (Fodor's version of) cognitive science.

McGinn's (1982, p. 229) analysis of meaning is that:

for *S* to mean that *p* is for *S* to be true iff *Q*, for some '*Q*' having the same truth conditions as '*p*', and for *S* to have some cognitive role  $\phi$  such that '*p*' also has cognitive role  $\phi$ .

McGinn's analysis specifies that *S* and *S'* have the same meaning only if they have identical conceptual roles. Field explicitly claimed that sameness of conceptual role is a necessary

condition for intro-person synonymy but explicitly denies its usefulness in characterizing inter-personal synonymy (Field, 1977). McGinn is unclear on this point, but if his analysis of ' $S$ ' means that ' $p$ ' is to be used by an interpreter to specify the meanings of a speaker's sentences, then it requires that the conceptual role of the interpreter's sentence ' $p$ ' and the conceptual role of ' $S$ ' be identical. However, only in science fiction thought experiments are the conceptual roles of Arabella's and Barbarella's sentences the same. As long as there is the slightest difference between Arabella's and Barbarella's probability assignments, no sentence will have the same conceptual role for both. If Arabella assigns a probability of 1, to 'it is raining' while Barbarella assigns it a probability of 0.2, then on McGinn's account their sentences have different meanings. Barbarella would be mistaken if she said that Arabella's sentence 'It is raining' means that it is raining. It is difficult to see how sense can be made of communication on this account of sameness of meaning. If Arabella and Barbarella assign different meanings to 'it is raining' (because they assign different probabilities), then there is nothing in common to be communicated. If they assign the same meaning (and so, have the same probability assignment), then there is no need for communication.

McGinn's account runs into similar difficulties as an account of 'means that' for internal representations. The usual view of advocates of mentalese accounts of belief (e.g., Fodor) is that 'Arabella believes that  $p$ ' is true iff she bears a certain relation  $R$  to a representation  $S$  which

means that  $p$ . Since it is improbable that any of Barbarella's sentences has the same conceptual role as  $S$ , Barbarella's claims concerning what Arabella believes are bound to be incorrect on McGinn's view.

The heart of the problem is that for  $S$  to mean that  $p$ ,  $S$  need not have precisely the same conceptual role as ' $p$ ' but rather a conceptual role appropriately similar to ' $p$ '. Until the DAT theorist has specified when the conceptual roles of  $S$  and ' $p$ ' are sufficiently similar to count as their having the same meaning (when they have the same truth condition) he has not adequately characterized ' $S$  means that  $p$ '. But this seems to be a hopeless task as long as conceptual role is characterized purely formally, e.g., in terms of probability relations.

Even if McGinn could overcome the difficulties just discussed, we will now show that the two components of McGinn's DAT, neither separately nor together, comprise a theory of interpretation. The adequacy conditions to be met by a DAT truth theory differ significantly from those to be met by a Davidsonian truth theory. McGinn views the truth theory as assigning 'facts' or 'states of affairs' to indicative sentences. This leads him to see a theory which issues in, e.g., "'Water is wet' is true iff H<sub>2</sub>O is wet' as adequate since, according to him, 'Water is wet' and 'H<sub>2</sub>O is wet' express the same fact. In fact, he says that any statements which are necessarily equivalent (substitutable in all non-psychological contexts) have the same truth conditions. Field's conception of the truth theory component is a bit different. He imagines the theory

issuing in theorems like: 'Water is wet' is true iff the stuff denoted by 'water' has the property denoted by 'is wet'. He hopes for a physicalistic theory of the denotation relation. It is clear that McGinn and Field see a truth theory as explicating the relation between language and the world. McGinn thinks that sentences are assigned facts. Field sees expressions as denoting bits of reality.

The first point to notice is that neither McGinn's, nor Field's, truth theories will serve as theories of interpretation (except in very unusual circumstances). We will discuss McGinn's account first. Suppose that Arabella utters 'Water is wet'. If Barbarella knows that 'Water is wet' is true iff  $H_2O$  is wet, then she will correctly conclude (following our first paradigm) that  $H_2O$  is wet. However, following our second paradigm, she may also conclude that Arabella believes that  $H_2O$  is wet, and this might very well be a mistake. The trouble is that even though 'Water is wet' is true iff  $H_2O$  is wet, this truth condition does not express the belief Arabella intends to communicate. Field's truth theory is even less adequate as a theory of understanding. If Barbarella knows that 'Water is wet' is true iff the stuff denoted by 'water' has the property denoted by 'is wet', then she will be able to conclude from Arabella's utterance that the stuff denoted by 'water' has the property denoted by 'is wet'. But there is a large gap between this and the conclusion that water is wet. It is certainly the latter, and not the former, that is communicated by Arabella's utterance.

The previous objections show that the truth theory component of a DAT does not provide a full specification of the information which a competent English speaker has, and brings to bear, in interpreting the utterances of others; so, it does not express the information known by a competent speaker. But, of course, Field and McGinn do not advance these theories as theories of understanding. Since the conceptual role component is supposed to characterize the mental aspects of meaning, perhaps it specifies information sufficient to interpret a speaker's utterances. So, we will now examine the conceptual role theory to see whether it will enable one to interpret another's language.<sup>4</sup>

We will first discuss McGinn's and Field's proposal concerning the representation of conceptual role, and then the less precise characterizations given by Sellars and Harman. We will suppose, once again, that Arabella utters 'Es schneit'. We also suppose that Cinderella knows the conceptual roles of all of Arabella's sentences; that is, she knows the conditional probability function  $P(A/B)$  defined on all the sentences of Arabella's language. Is this information sufficient to enable her to justifiably infer either that it is snowing or that Arabella believes that it is snowing? The answer is 'No!' Suppose that Cinderella starts by knowing that Arabella believes 'Es schneit' is true (or perhaps that she assigns a probability close to 1 to 'Es schneit'). How can the conceptual role theory support Cinderella's inference to the belief that it is snowing? It may support Barbarella's concluding that Arabella assigns a high probability to

'Es ist kalt', a low probability to 'Das Wetter ist schön', and so on. She might even be able to predict how Arabella would behave, what sounds she will utter next, etc., but she will not know what Arabella means. In particular, she will not be in any position to infer that it is snowing or that Arabella believes that it is snowing. Unless Barbarella knows that 'Es schneit' is true iff it is snowing, Arabella's assigning a probability of 1 to 'Es schneit' does not count as evidence that it is snowing. Nor does it count as evidence for Arabella believing that it is snowing. Knowledge of Arabella's probability assignments is not sufficient for interpreting her utterances, at least not if that understanding includes the capacity to infer that it is snowing, that Arabella believes that it is snowing, and so on.

One might think that the problem with Field's conceptual role semantics (as a characterization of knowledge sufficient for language understanding) is that it fails to characterize relations between expressions and the world. In contrast, Sellars and Harman sketch a conceptual role theory which includes causal relations between sentences and the world. Harman says:

There is no suggestion that content depends only on functional relations among thoughts and concepts, such as the role a particular concept plays in inference. Of primary importance are functional relations to the external world in connection with perception, on the one hand, and

action on the other.

Harman does not specify the structure of a theory which characterizes conceptual role non-solipsistically. Lycan (1984, Ch. 10) in the course of criticizing Harman's proposal, suggests that such a theory would imply statements like:

(A) The usual cause of Arabella's uttering 'Es schneit' is her being in belief state *K*, and the usual cause of her being in belief state *K* is it's snowing in the vicinity.

If Barbarella knew this, then she would be in a position to infer that it's snowing from Arabella's having uttered 'Es schneit'. This would be a case of inference to the best explanation, rather than an instance of our first paradigm. Truth seems to play no special role. As Lycan is well aware, there are a number of problems with this proposal. It is not plausible for most sentences, especially theoretical ones. The usual cause (if there is such a thing) of my uttering 'There are positively charged electrons' is not my being in a belief state which is itself caused by there being positively charged electrons. The latter state of affairs may figure in the causal history of my utterance, but it is only one among many causes and not the 'usual cause'. If we put this objection aside, (A) could be used to draw conclusions about the world from Arabella's

utterances. But it does not enable one to draw appropriate conclusions about what Arabella believes when she utters, e.g., 'Water is wet'. The problem is that if water's being wet is the cause of her utterance, so is H<sub>2</sub>O's being wet. In fact, Lycan's proposal adds nothing more to the theory of interpretation than the truth theory component provides.

We can imagine McGinn replying to our discussion with the suggestion that although neither component of a DAT by itself is a theory of interpretation, the combination is. One might be encouraged to think this since, according to McGinn, the conjunction of the two theories and his definition of meaning (M) yield theorems of the form '*S* means that *p*', for each sentence *S* of *L*. Since '*S* means that *p*' entails '*S* is true iff *p*', we could employ these theorems in our two paradigms. But a closer look at McGinn's analysis of '*S* means that *p*' reveals that the two component theories do not entail statements of the form '*S* means that *p*'. Recall McGinn's definition:

(M) *S* means that *p* iff *S* is true iff *Q*, for some '*Q*' having the same truth conditions as '*p*', and for *S* to have some cognitive role  $\phi$  such that '*p*' also has cognitive role  $\phi$ .

The truth theory component may entail, e.g., 'Wasser ist naß' is true iff H<sub>2</sub>O is wet. The conceptual role component may entail 'Wasser ist naß' has the same conceptual role as 'Water is

wet' (and a different conceptual role from 'H<sub>2</sub>O is wet'). But we can put the two components together to explain that 'Wasser ist naß' means that water is wet only if we know that 'Wasser is naß' means that water is wet. Of course, we knew that if we knew English. But McGinn cannot rely on this since a theory of interpretation is supposed to characterize knowledge sufficient for understanding a language without presupposing the understanding of any language. The difficulty here is the same as we encountered in considering the conceptual role component by itself. Knowing that *S* has the same conceptual role as '*p*' is not sufficient for understanding *S* (Twin-Earth problems aside) unless one knows what '*p*' means. Contrast this with knowing that *S* is true iff *p*. Someone might have this knowledge (the knowledge expressed by '*S* is true iff *p*') without knowing what '*p*' means. We want a theory of interpretation to state information sufficient for interpretation which itself does not depend on understanding the language in which it is formulated. It is this that (M) fails to do. Our conclusion is that the Field-McGinn conceptual role theory is inadequate as a theory of interpretation.

What is the relation between McGinn's two component theory and a Davidsonian truth theory? According to McGinn (1982, p. 240)

...the Davidsonian perspective, while not actually being incorrect - for it is, after all, tacitly a dual component conception - is apt to deceive us about the theoretical resources we need in a full

adequate theory of meaning.

McGinn thinks that a Davidsonian theory is 'tacitly a dual component conception', since in a Davidsonian T-sentence '*S* is true iff *p*', '*p*' not only specifies a truth condition of *S* but is also supposed to be a translation of *S*. If translation is understood as requiring identity (or rather, isomorphism) of conceptual role, then it seems to McGinn that a Davidsonian theory is simply a misleading formulation of a two component theory.

McGinn is correct in thinking that a Davidsonian theory addresses two aspects of meaning. It accounts for language world relations by characterizing truth and reference, and it accounts for mental aspects by characterizing knowledge sufficient for interpreting a language. But we strongly disagree with McGinn's claim that Davidson's theory is tacitly a dual component view presented in a misleading manner. There are significant differences between a Davidsonian theory and McGinn's DAT. First, a Davidsonian theory does not require for the adequacy of its T-sentence theorems that *S* and '*p*' have the same conceptual role. We have already discussed the implausibility of this requirement. Davidson's theory is subject to the much looser constraint that it leads to reasonable attributions of beliefs and other propositional attitudes to speakers. Second, as we have just argued, McGinn's DAT is not a theory of interpretation. Knowledge of it does not enable one to determine what Arabella believes on the

basis of her sincere utterances. This is the heart of our objection to McGinn's and other two-component theories. Whatever insight, if any, may be gained from decomposing meaning into two separate features, the resulting characterizations are not suitable as a theory of interpretation. Knowledge sufficient for understanding a language cannot be extracted and separated from knowledge of reference and truth since to understand a sentence is to know what it says about the world.

We have yet to consider the question of whether DATs provide adequate accounts of meaning for mental representations. It is *prima facie* plausible that DATs are just what is needed for mentalese. Recall that Fodor argues that cognitive science requires a methodologically solipsistic notion of content in its explanations of behavior. It is not unreasonable to look to the conceptual role component of a DAT to characterize this notion of content, while the truth theory component characterizes a broader notion of content involving relations between mental representations and the world. Since mentalese is not a language in which anyone communicates, our arguments that DATs do not provide theories of interpretation may appear irrelevant. In fact, some authors explicitly endorse dual aspect accounts for mentalese, but not for public languages (Lycan,1985).

In the course of our discussion of Fodor in the first section, we saw that Fodor thinks that cognitive science requires a solipsistic notion of content. That is, a characterization of

content which conforms to the formality condition: no difference in content without a difference in formal properties. The Field-McGinn characterization of conceptual role is solipsistic. But, we will argue, it does not yield a specification of content that is methodologically solipsistic, for the simple reason that *it yields no specification of content at all*. We have, in effect, already established this in our discussion of DATs as theories of interpretation. The problem is that a complete specification of the conceptual roles of the sentences *S* of Arabella's language of thought does not enable us to fill in the blanks in 'If Arabella bears *R\** to *S*, then she believes that -----'. But this is what Fodor requires of a characterization of the contents of mental representations.

It might be thought that the dispute between us and someone who holds that conceptual role provides a solipsistic notion of content is merely a semantic quibble about what is to count as content. But it is not, since the operative notion of content is intended (by Fodor, who formulated this problem) to play a particular role in psychological explanations. Consider the following 'psychological' explanation. Arabella jumped because she believed that by jumping she would cause it to rain and she wanted it to rain. The phrases following 'believed that' and 'wanted' express contents. They explain Arabella's action by citing causes (her beliefs and desires) which 'rationalize' it. If we describe the causes of her actions without citing the contents of the propositional attitudes, then the resulting explanation no longer has its 'rationalizing'

force. An explanation of Arabella's jumping which employs conceptual role would go something like this. Arabella bears  $B^*$  to a mental representation which has conceptual role  $Q$  and  $W^*$  to a mental representation which has conceptual role  $V$ ; so, she jumped. Presumably the conceptual roles would be characterized in such a way that jumping normally follows upon bearing these relations to those representations. It is clear that the explanatory force of this explanation, whatever its value, does not rationalize Arabella's behavior. Only an explanation which appeals to content can do that.

#### DAVIDSONIAN TRUTH THEORIES SOLVE THE PROBLEMS

We have shown that McGinn's two component definition of meaning is inadequate, that his DAT is not suitable for use as a theory of interpretation, and that it fails to yield a solipsistic notion of content for mentalese. Also, we began to show how a Davidsonian truth theory, by providing a unified account of understanding and reference, is suitable for use as a theory of interpretation. Still, we need to inquire into the extent to which Davidsonian theories can cope with the problems which motivated the DAT proposal. So, we will now consider whether a Davidsonian truth theory can be used to interpret sentences containing proper names, indexicals, and natural kind terms. We will conclude with a few remarks on the prospect of obtaining a methodologically solipsistic theory of content.

What T-sentences will a Davidsonian truth theory entail for a language with proper

names? It is clear that, in contrast to the truth theory component of a DAT, a Davidsonian theory must be able to assign different truth conditions to, e.g., 'Cicero is bald' and 'Tully is bald'. DAT truth theories maintain that these sentences have the same truth conditions, since it is metaphysically necessary that they are equivalent (assuming that names are rigid). If one thinks, as McGinn does, that a truth theory assigns possible states of affairs of facts to indicative sentences, then we can see why DAT truth theories assign the same truth conditions to the two sentences. But it is not necessary to think of truth theories in this way. It is clear that Davidson rejects the reification of truth conditions as states of affairs (Davidson, 1969).

There is no reason why a Davidsonian truth theory cannot contain (1) and (2) without containing (3) and (4) (McDowell, 1980):

1. 'Cicero is bald' is true iff Cicero is bald.
2. 'Tully is bald' is true iff Tully is bald.
3. 'Cicero is bald' is true iff Tully is bald.
4. 'Tully is bald' is true iff Cicero is bald.

Arabella might believe (1) and (2) without believing (3) and (4). In that case it would be a mistake to include (3) and (4) in a theory intended to be employed in interpreting her utterances.

If (3) and (4) are included in a truth theory for Arabella's language, then one would be licensed to infer (via paradigm II) that Arabella believes that Tully is bald from the fact that she utters 'Cicero is bald'. Suppose that a truth theory *T* contains (1) and (2), but not (3) and (4), and that Arabella assents to 'Cicero is bald' and dissents from 'Tully is bald'. We can use *T* to conclude that Arabella believes that Cicero is bald and believes that Tully is not bald. But there is no contradiction forthcoming unless we have some principle in the theory that permits the substitution of coreferential names in belief contexts. That it leads to attributing contradictory beliefs to Arabella is sufficient reason for rejecting the principle.

Our account of the truth conditions of sentences containing names is compatible with the view that names rigidly designate. If names are rigid designators, then true identity statements composed of names are necessarily true. Suppose that we add to *T* the 'axiom' - It is necessary that (Cicero = Tully). We can also suppose that Arabella knows that if Tully = Cicero then it is necessary that (Tully = Cicero). Of course, we will not add the axiom that Arabella knows that it is necessary that (Cicero = Tully), since that is false. Even with these additions, we cannot derive from *T*, and the fact that Arabella assents to 'Cicero is bald' and dissents from 'Tully is bald', that she believes that Cicero is bald and that Cicero is not bald. Of course, if one thinks that co-referential rigid designators are substitutable *salva veritate* in belief contexts, then the consequence will be to saddle Arabella with contradictory beliefs. As

we mentioned earlier, that seems to us to be reason to reject substitutivity.

A defender of DATs is likely to respond to our proposal in the following fashion. A Davidsonian truth theory is supposed to be an empirical theory, but what evidence could distinguish between a theory which contains (1) and (2), and a theory which contains (3) and (4)? Our reply, of course, is that the two theories are empirically discriminated in their applications in paradigm II. Assuming that Arabella might believe that Tully is bald without believing that Cicero is bald (and that we can obtain evidence for this), we can distinguish between the two theories. What evidence is relevant to whether Arabella believes that Tully is bald or believes that Tully is not bald? The dual aspects theorist's answer is that conceptual role is relevant. If Arabella bears B\* to an internal representation 'Tully is bald' which has a conceptual role similar to the conceptual role of my sentence 'Tully is bald' then she believes that Tully is bald. But, as we argued previously, identity, or similarity, of conceptual role is too strong a requirement to make us co-believers. Conceptual role seems especially irrelevant when it comes to translating proper names. What does seem relevant is the history of acquisition of a name. Suppose, for example, that I can trace Arabella's current use of 'Cicero' and my current use of 'Cicero' back to a common source, say, we both acquired the name upon hearing it spoken in a history class. Then that would count in favor of my attributing to her the belief that Cicero is bald upon hearing her utter 'Cicero is bald' even if the conceptual role of her sentence

is quite different from mine. Of course, considerations governing the interpretation of names are quite complex. Our point is that whatever they are (a) conceptual role is not of great importance, and (b) a Davidsonian truth theorist can (indeed should) avail himself of these considerations in fashioning a truth theory.

Indexicals provide another motivation for DATs. Arabella's and twin-Arabella's understanding of 'I am 30 years old' is the same, but the truth conditions of their utterances of the sentence differ. DATs say that the conceptual roles of the sentences are the same (or isomorphic) while the truth conditions of their utterances differ. As we observed in the second section, a conceptual role theory does not seem to provide an adequate account of our understanding of indexicals. One's understanding of, for example, 'I am now in California', involves knowing how to infer what was said from knowledge of the context in which the sentence is uttered. It is not clear that this can be represented within Field's framework for conceptual role. Be that as it may, we now want to show how truth theories can be modified to interpret sentences with indexicals.

A number of authors have proposed ways of extending truth theories to languages with indexicals (Davidson, 1972; Taylor, 1980). However, no one seems to have addressed the problem of showing how a truth theory for languages with indexicals can be employed to interpret the utterances of a speaker of the language. It may be thought that a truth conditional

account will not work for indexicals because 'Arabella is tall' and 'I am tall' uttered by Arabella have the same truth condition, namely, that Arabella is tall. But clearly Arabella's understanding of the two sentences might be different, since she might believe what one expressed and not what the other expressed. DATs deal with this by claiming that there is something other than truth conditions involved in understanding. What we are going to do is to show how a truth conditional account of indexicals can yield an account of understanding according to which Arabella and Barbarella have the same understanding of, e.g., 'I am in California', and yet different information can be communicated by their utterances of the sentence. Our idea builds on Davidson's suggestion that the truth predicate applies to a sentence at a time, for an utterer, at a place (with perhaps further relativizations required).

On our view indexical sentences possess *general truth conditions*. For example, understanding 'I am now reading Russell' includes knowing that the following is common knowledge among speakers of English:

(C)  $(x)(t)$  ('I am now reading Russell' is true for x at t iff x is reading Russell at t)

Interpreting a language with indexicals involves introducing a relativized truth predicate. With 'I' and 'now' as the only indexicals in this sentence, the relativized truth predicate is 'true for x at

t', where Y ranges over utterers and 't' over times. How can knowledge that (C) enable someone to interpret utterances? Suppose that Arabella says to Barbarella 'I am now reading Russell'. Then Barbarella may be able to reason as follows:

(1). Arabella uttered 'I am now reading Russell' at noon.

So,

(2). 'I am now reading Russell' is true for Arabella at noon.

(C)  $(x)(t)$  ('I am now reading Russell' is true for x at t iff x is reading Russell at t).

So,

(3). Arabella is reading Russell at noon.

Barbarella is justified, in part, in believing that Arabella is reading Russell at noon on the basis of her hearing Arabella's utterance because she knows that (C). But notice that she also makes use of information about the context of utterance. In particular, she uses her belief that Arabella is the utterer and that she made her remark at noon. If Barbarella failed to believe this, she would not be in a position to conclude (3). Exactly what she can conclude depends on what she believes. If she believes that the utterer is the tallest woman in the room and that the time is the same time that Cinderella laughed, then Barbarella could employ (C) to learn from

Arabella's remark that the tallest woman in the room was reading Russell at the same time that Cinderella laughed.

Of course, there are ways that any competent interpreter has of identifying an utterer and the time of utterance. For example, if Barbarella hears someone utter 'I am now reading Russell', even if she knew nothing else about the utterer, she would know that he is the utterer of that utterance. She could then use (C) to conclude that that utterer of 'I am now reading Russell' is reading Russell.

We can also explain how Barbarella can employ her knowledge that (C) is common knowledge among speakers of English to learn something about what Arabella believes when she utters 'I am now reading Russell'. But here the inferences are a bit more delicate. Considering the following reasoning:

- (1). Arabella utters 'I am now reading Russell' at noon.
- (4). Arabella believes that 'I am now reading Russell' is true for Arabella at noon.
- (5). Arabella believes that  $(x)(t)$  ('I am reading Russell' is true for  $x$  at  $t$  iff  $x$  is reading Russell at  $t$ ).

So,

- (6). Arabella believes that Arabella is reading Russell at noon.

In this bit of reasoning Barbarella employs information concerning Arabella's beliefs about whom she is and about when her utterance occurs. To conclude (6), she uses the information that Arabella believes that Arabella produced the utterance and that Arabella believes that it occurs at noon. If Arabella did not know that she is Arabella or did not know that the time of the utterance is noon, then Barbarella would be mistaken in concluding that (6).

If Barbarella does not know any singular term (or a translation thereof) that she believes Arabella believes refers to herself, then she will be unable to employ the knowledge embodied in the general truth condition (C) to arrive at conclusions about when Arabella believes herself to be reading Russell at noon. Again, there are some ways that any competent speaker of English has of identifying herself and the time of her utterance. Arabella will always know (if she understands English) that 'I' refers to herself. Of course, Barbarella cannot use 'I' to refer to Arabella, but she can express what Arabella believes by ascribing to her the belief that the utterer of that utterance believes of herself that she is reading Russell, from which she can conclude that Arabella believes that she (herself) is reading Russell.

Knowledge that (C) together with other information (some available to every competent speaker, some not) enables an interpreter to learn both about the utterer's beliefs and about the world beyond the utterer. Suppose that, as in Perry's story, Hume and Heimson each utter 'I am

a Scottish philosopher' (Perry, 1979). Each believes himself to be Hume. If we know all this, we can employ the general truth condition - (x) ('I am a Scottish philosopher' is true for x iff x is a Scottish philosopher) - to conclude from Hume's utterance that Hume is a Scottish philosopher and that he believes himself to be a Scottish philosopher. Since we know that Heimson is deluded, we would not conclude from his utterance that he is a Scottish philosopher. But we can conclude, following the second inference pattern above, that Heimson believes himself to be a Scottish philosopher.

We have indicated how general truth conditions can play a role in a theory of interpretation. To provide an adequate account we would at least have to show how to construct a truth theory for a language with indexicals, which has correct general truth conditions as theorems and we would need to develop a logic in which the inferences to conclusions about the world and the utterer's beliefs can be represented. We leave these tasks to a future paper.

Natural kind terms presented another problem which motivated the development of DATs. Can a truth theory of interpretation be constructed for languages containing natural kind terms? Imagine an English speaking radical interpreter who arrives on Twin-Earth. He notes the patterns of sentences held true (and the degrees to which they are held true) and begins to devise a theory of interpretation. Suppose that he is unaware that the stuff they call 'water' is composed of XYZ molecules. Then he is likely to employ the homophonic truth theory, which contains the

clause "'Water is wet' is true (as uttered by a Twin-Earther) iff water is wet'. When twin-Arabella utters 'This is water', interpreter will conclude that this is water and that twin-Arabella believes that this is water, just as he would conclude if he were interpreting Arabella on Earth. But he will be mistaken. This is not water (since it is composed of XYZ) and twin-Arabella does not believe that it is water (since she has never encountered water). However, these mistakes will pass undetected as long as the interpreter does not know the compositions of water and twin-water. Once he discovers that the chemical composition of water is H<sub>2</sub>O and the composition of twin-water is XYZ, he will notice that he sometimes incorrectly interprets twin-Arabella. How might he revise his theory in the face of these incorrect interpretations?

He could try "'Water is wet' is true in twin-English iff XYZ is wet'. This change will lead him to correctly conclude that this is XYZ when he hears twin-Arabella utter 'This is water', but it might also lead him to mistakenly conclude that twin-Arabella believes that this is XYZ. This would be a mistake if twin-Arabella did not know the chemical composition of the stuff she calls 'water'.

The solution is for interpreter to enrich his own language by adding a term 'water' with the stipulation that it rigidly designates XYZ, but has the same conceptual role as Arabella's word 'water'. His truth theory for twin-English should contain the theorem "'This is water' is true in twin-English iff this is water'.

We have shown how understanding of a language which contains names, indexicals, and natural kind terms, can be characterized by the knowledge of a truth theory which satisfies certain constraints. We argued that the DAT account is not successful, since the conceptual role component cannot explain how we acquire information from the utterances of others. However, it is clear that our Davidsonian account of understanding does not satisfy one of the constraints that the conceptual role component was designed to satisfy; on the Davidsonian account, meaning is not entirely in the head. Arabella and twin-Arabella may be physiological identical, but the truth theories that describe their understanding of their languages are different, as we saw in interpreting natural kind terms in English and twin-English. We have argued elsewhere (Lepore and Loewer, 1986) that this does not show Davidsonian truth theories to be inadequate, but rather that the information that an interpreter can use does not supervene on physiology. The view that it is possible to describe a person's understanding of a language at a level at which understanding supervenes on physiology and which is also able to account for communication is wrong.

## **CONCLUSION**

Dual Aspect theories claim that there are two aspects to meaning: one which relates language to the world and another which relates language to the mind. McGinn goes on to claim that the most perspicacious theory of meaning is one which gives distinct and independent treatments of

each aspect: a theory of truth and a theory of conceptual role. We agree that meaning has two aspects but we disagree concerning the appropriate form for a theory of meaning. In this paper we showed that McGinn's dual component theory is incapable of serving as a theory of interpretation and for that reason is not the most perspicacious form for a theory of meaning. McGinn confuses the evidential basis for theory of meaning which may include information about the causal relations between representations and the world and information about conceptual role with a theory of meaning. In contrast, Davidsonian truth theories can accommodate both the aspects while providing a theory of interpretation.

#### **ACKNOWLEDGMENTS**

We would like to thank Ned Block, Jerry Fodor, Brian Maclaughlin, Richard Grandy, Gilbert Harman, Bernard Linsky, John McDowell, Colin McGinn, Hilary Putnam, and Stephen Schiffer, for comments on, and discussion of, earlier drafts of this paper. We especially would like to thank John Biro, Donald Davidson, and William Lycan for their help. Earlier drafts of this paper were read at the University of Michigan, Central Michigan University, Rice University, The Pacific APA, 1984, the Florence Center for the Philosophy and History of Science, the University of Urbino, and the City University of New York.

#### **NOTES**

1. Except for XYZ replacing H<sub>2</sub>O in twin-Arabella's body, but this, of course, is irrelevant to the

argument. Putnam distinguishes 'narrow psychological states' from 'broad psychological states'. Only the former supervene on brain states described in neurophysiological terms. Clearly Putnam supposes that 'grasping a sense' is a narrow psychological state.

2. McGinn envisions a probability assignment defined over a language which contains indexical sentences. But there are problems in interpreting probabilities of an indexical sentence. One difficulty is accounting for changes in the probability of, e.g., 'The meeting begins now', as time passes. We also note that a probability assignment does not capture the indexicality of indexical sentences, i.e., the way truth value depends on context.

3. We switch languages here from English to German only to emphasize the informational value of T sentences.

4. It is obvious that the truth theory component of a DAT does not assign content to sentences of mentalese in a way that respects the formality condition. DAT assignments of truth conditions does not respect our ordinary individuation criteria for belief attributions. A representation R might be assigned either the truth condition that water is wet, or equivalently, from the perspective of a DAT, that H<sub>2</sub>O is wet. While there may be a 'transparent'

sense of belief for which the belief that water is wet is the same as the belief that H<sub>2</sub>O is wet, it is clear that this assignment of belief contents is inappropriate for CTM.

## REFERENCES

- Block, N.1986. 'Advertisement for a Semantics for Psychology.' Midwest Studies in Philosophy. Studies in the *Philosophy of Mind*. French, P.A. et. al. (eds.) Minneapolis: University of Minnesota Press.
- Davidson, D. 1984. *Inquiries Into Truth and Interpretation*. Oxford: Oxford University Press.
- Davidson, D.1967. *Truth and Meaning*. In Davidson (1984).
- 1968. *On Saying That*. In Davidson (1984).
- 1969. *True to the Facts*. In Davidson (1984).
- 1973a. *Radical Interpretation*. In Davidson (1984).
- 1973b. In *Defense of Convention T*. In Davidson (1984).
- 1974. *Belief and the Basis of Meaning*. In Davidson (1984).
- Field, H.1977. 'Logic, Meaning and Conceptual Role.' *Journal of Philosophy* 7.
- 1978. 'Mental Representation.' *Erkenntnis* 13.
- Fodor, J.1975. *The Language of Thought*. Cambridge: Crowell.
- 1980. 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology.' *Behavioral and Brain Sciences* 3.

- 1982. 'Cognitive Science and the Twin-Earth Problem.' *Notre Dame Journal of Formal Logic* 23.
- Narrow *Content*. *Unpublished* manuscript.
- Harman, G. 1973. *Thought*. NJ: Princeton University Press.
- 1974. 'Meaning and Semantics.' In: M.R. Munitz and P.K. Unger (Eds.), *Semantics and Philosophy*. New York: New York University Press.
- 1982. 'Conceptual Role Semantics.' *Notre Dame Journal of Formal Logic* 20.
- 1987. Non-Solipsistic Conceptual Role Semantics. In: Lepore and Loewer (eds) *New Directions in Semantics*. London: Academic Press Inc. Ltd.
- Jeffrey, R. 1983. The First *Definetti Lecture*. *Unpublished* manuscript.
- Kaplan, D. 1977. *Demonstratives*. *Unpublished* manuscript.
- Kripke, S. 1972. *Naming and Necessity*. In D. Davidson & G. Harman (Eds), *Semantics of Natural Languages*. Dordrecht: Reidel.
- 1979. A Puzzle about Belief. In Margalit (ed), *Meaning and Use*. Dordrecht: Reidel.
- LePore, E. 1982a. 'Truth and Inference.' *Erkenntnis* 18.
- 1982b. 'In Defense of Davidson.' *Linguistics and Philosophy* 5.
- 1983. 'What Theoretical Semantics Cannot Do.' *Synthese* 54.

**E. LEPORE AND B. LOEWER**

-----and Loewer, B.1981. 'Translational Semantics.' *Synthese* 48.

----- and ----- 1983. 'Three Trivial Truth Theories.' *Canadian Journal of Philosophy*.

----- and -----1986. 'Solipsistic Semantics.' *Midwest Studies in Philosophy*.

Loar, J.1981. *Mind and Meaning*. Cambridge: Cambridge University Press.

Loewer, B. 1982. 'The Role of Conceptual Role Semantics.' *Notre Dame Journal of Formal Logic* 23.

Lycan, W.1981. 'Form, Function and Feel.' *Journal of Philosophy* 5.

-----1982. 'Toward a Homuncular Theory of Believing.' *Cognition and Brain Theory* 4.

-----1984. *Logical Form*. Cambridge: M.I.T. Press.

-----1985. 'Paradox of Naming'. In B.K. Matilal & J.L. Shaw (Eds), *Analytical Philosophy Comparative Perspective*. Dordrecht: Reidel.

McDowell, J. 1980. 'On the Sense and Reference of a Proper Name.' In M. Platts (Ed.), *Reference, Truth and Reality*. London: Routledge & Keagan Paul.

McGinn, C. 1982. 'The Structure of Content.' In A. Woodfield, *Thought and Object*.

Perry, J.1978. 'Frege on Demonstratives.' *Philosophical Review* 59.

- 1979. 'The Problem of the Essential Indexical.' *Nous* 12.
- Putnam, H. 1975. 'The Meaning of "Meaning".' *Mind, Language and Reality: Philosophical Papers*. Cambridge: Cambridge University Press, vol. 2.
- 1986. *Meaning Holism*. In L. Hahn and P. Schilpp (eds.), *The Philosophy of W.V. Quine: The Library of Living Philosophers Vol. XVIII*. La Salle, Illinois: Open Court
- Computations and Interpretations. (CI)* Unpublished manuscript.
- Sellars, W. 1956. 'Empiricism and the Philosophy of Mind. *Minnesota Studies in the Philosophy of Science*. Minnesota: Minnesota Press, Vol. 1.
- 1963. 'Some Reflexions on Language Games.' *Science, Perception and Reality*.
- 1969. 'Language as Thought and as Communication.' *Philosophy and Phenomenological Research*, 29.
- Tarski, A. 1956. 'The concept of Truth in Formalized Languages.' *Logic, Semantics, and Metamathematics*. Oxford: Clarendon Press.
- Taylor, B. 1980. 'Truth Theories for Indexical Languages.' In M. Platts (Ed.), *Reference, Truth and Reality*. London: Routledge and Keagan Paul.
- White, S. 1982. 'Partial Character and the Language of Thought.' *Pacific Philosophical Quarterly*, 63.
- Zemach, E. 1976. 'Putnam and Meaning and Reference.' *Journal of Philosophy*.

\*'Dual Aspect Semantics', New *Directions in Semantics*, E. LePore (ed.) Academic Press Inc.;  
reproduced here by kind permission of Academic Press Inc. and E. LePore and B. Loewer.

*S. Silvers (ed.), Rerepresentation, 161-188. ©1989 by Kluwer Academic Publishers.*