

The Role of "Conceptual Role Semantics"

BARRY LOEWER*

In his paper Gilbert Harman defends conceptual role semantics (CRS), a theory of meaning he has been elaborating for the past decade ([11]-[13]). CRS is especially interesting for the way it combines issues that are central to both the philosophy of language and the philosophy of mind. According to Harman it is founded on two claims:

- I. The meanings of linguistic expressions are determined by the contents of the concepts and thoughts they can be used to express.
- II. The contents of concepts and thoughts are determined by their "functional role" in a person's psychology.

A corollary of I and II is that the use of symbols in calculation and thought is more basic than the use of symbols in communication. It seems to me that a good way to appreciate the role of CRS is to compare it with truth conditional semantics (TCS). Harman seems to consider the two approaches to be to some extent in competition with each other and has argued that TCS can make at most a subsidiary contribution to the theory of meaning. In my comments I will argue that the two approaches are best seen as complementary. Although CRS is a significant contribution, TCS has a central role to play both in accounts of language used for communication and in language used for calculation and thought. In Section 1 I will sketch an account of understanding language used for communication which is based on TCS. In Section 2 I consider an argument of Harman's which he thinks demonstrates the impotence of TCS. In Section 3 I consider some contributions of CRS and in 4 I argue that TCS is central to an account of language used in thought.

*I would like to thank Robert Laddaga and Ernie Lepore for helpful discussions on the issues surrounding conceptual role semantics.

1 Concerning the view of Davidson ([1] and [2]), Lewis [16], and others that an account of the truth conditions of the sentences of a language is central to a theory of meaning for that language, Harman remarks

... this seems wrong. Of course if you know the meaning in your language of the sentence *S*, and you know what the word "true" means, then you also know something of the form "*S* is true iff" ... But this is a trivial point about the meaning of "true" not a deep point about meaning.

Harman goes on to grant that a theory of truth for a language may shed light on meaning but only by specifying implications among sentences. His view is that the functional roles of logical constants 'and', 'or', 'all', etc., are mostly characterized by specifying their roles in inference. So a theory of truth can at best specify the meanings of logical constants and does so by characterizing their conceptual roles.

I think that Harman's dismissal of truth conditional semantics is mistaken. I first want to consider an argument which shows that TCS has an important role to play in an account of communication and language understanding.

One of the goals of a theory of meaning is to characterize linguistic competence. This problem has been construed by Dummett [6] and Davidson as the problem of characterizing what someone must know to understand a language. Davidson's answer is, of course, that understanding a language consists, at least in part, in knowing the truth conditions of the sentences of the language. There is a simple argument, suggested by Davidson but never explicitly formulated by him, which seems to show that his answer is correct. (This argument is spelled out in more detail in [15].) Consider the following communication episode. Arabella, Barbarella, and Esa are in a room with Arabella looking out the window. Arabella and Barbarella understand German but Esa does not. Arabella turns from the window to Barbarella and Esa and utters the words "Es schneit". On the basis of this utterance Barbarella comes to believe that it's snowing (and also that Arabella believes that it's snowing, etc.) while Esa comes to believe only that Arabella said something which is probably true. We can focus on the question of what knowledge comprises Barbarella's understanding "Es schneit" by asking what would Esa need to know to come to the same beliefs as Barbarella. The obvious candidate for this knowledge is the knowledge that "Es schneit" is true iff it's snowing.¹ A reconstruction of the reasoning which justifies Barbarella's acquisition of the belief that it's snowing looks like this:

1. Arabella utters the words "Es schneit"
2. Since "Es schneit" is an indicative sentence and since Arabella is generally reliable, her utterance of "Es schneit" is true
3. "Es schneit" is true iff it's snowing

therefore

4. It's snowing.

Both Esa and Barbarella can come to believe that "Es schneit" is true by knowing a bit of German grammar (enough to recognize indicative sentences) and knowing that Arabella is reliable. But only Barbarella is in a position to go

on to conclude that it's snowing since only she understands German. And if my argument is correct that understanding must consist in part in knowing the truth conditions of the German sentence.

An objection to my argument is that understanding a language does not involve propositional knowledge, knowledge that. Some, including Harman [13], have argued that understanding is a know how which cannot be explicated in terms of propositional knowledge. I do not want to take up this thorny issue here except to make one observation. Even if in our ordinary understanding of a language used for communication we do not employ propositional knowledge, I think it is clear from my example that anyone who understands such a language does know the truth conditions of its sentences.

Since there are infinitely many sentences of a language, a theory is needed to specify the knowledge in which understanding consists. As Davidson has been urging for over a decade, the kind of theory that is required is a Tarski-type theory of truth since it assigns truth conditions to each indicative sentence of a language. A theory of truth will, as Harman notes, also spell out implications among sentences, but more importantly, it will provide the core of an account of the understanding of language used in communication.

2 Harman is unimpressed by truth conditional semantics. I think his reasons are half based on a confusion and half based on an argument which he thinks shows that TCS merely postpones the problem of providing semantics. The argument also shows why he thinks language used in thought is more basic than language used in communication. First, the confusion. Harman says that "'Snow is white' is true iff snow is white" expresses a trivial point about the meaning of "true" not a deep point about the meaning of 'Snow is white'. I suspect that the reason he thinks this is that anyone who is familiar with quotations and the disquotational effect of "is true" can recognize that "'Snow is white' is true iff snow is white" expresses a truth, even if they do not have the slightest idea what 'Snow is white' means. This is correct but irrelevant. A theory of truth employed as an account of understanding does not attribute knowledge that the sentence "'Snow is white' is true iff snow is white" is true to a speaker but the knowledge that "Snow is white" is true iff snow is white. There is a world of difference between the two knowledge attributions. The second but not the first justifies the inference from a belief that the utterance "Snow is white" is true to a belief that snow is white. It is very easy to make the mistake of thinking that sentences like "'Snow is white' is true iff snow is white" are trivial when the single quoted sentence is in a language you understand. But the reason is that you exploit your understanding of the language to recognize their truth. The truth conditional semanticist is in fact claiming that your understanding "Snow is white" partially consists in your knowing its truth conditions.

The argument that underlies Harman's dismissal of truth conditional semantics is formulated in another paper:

Davidson would [presumably] say that the speaker understands [the sentence 'Snow is white'] by virtue of the fact that he knows it is true if and only if snow is white. The difficulty . . . is that [for the speaker to know this he] needs some way to represent to himself snow's being white. If the relevant speaker

uses the words 'Snow is white' to represent in the relevant way that snow is white . . . Davidson's (theory) would be circular. And, if speakers have available a form of Mentalese in which they can represent that snow is white, so the (theory avoids) circularity, there is still the problem of meaning for Mentalese. ([12], p. 286)

Harman's argument can be reconstructed as follows:

1. Suppose that to understand 'Snow is white' is to know that 'snow is white' is true iff snow is white.
2. But to know that *S* one must have a way of representing *S* to oneself.
3. That is to know that *S* there must be in one's language of thought a token which means that *S*.
4. If the language of thought is English then Davidson's theory is circular.
5. If the language of thought is Mentalese then Davidson's theory is incomplete since it fails to specify semantics for Mentalese.

Harman seems to take this argument as establishing his view that TCS makes no contribution (other than helping to characterize the functional role of certain words) to the theory of meaning. But so strong a conclusion is not supported by the argument. I have previously argued that understanding a language used for communication involves knowing the truth conditions of its sentences and nothing in Harman's argument undermines this claim. Neither does Harman's argument show that sentences in a language of thought, if there is such a language, do not have truth conditions or that characterizing their truth conditions is not an important part of characterizing their meanings. What his argument does show is that if we grant the truth of 3 then understanding a language of thought cannot be explained as *knowing* the truth conditions of its sentences. To attempt to do so would involve one in a regress since we would then have to explain knowledge of the truth conditions of a language of thought by postulating another language of thought understood by the knower. This is an important point. It does show that if 3 is correct then TCS is not the whole story concerning meaning. (Similar points are made by Field in [8].)

Premise 3 of the argument is crucial. Sometimes the view that to believe that *p* is to be related to a token of some language of thought (English or Mentalese) which means that *p* seems to be taken to be a conceptual truth. But this is a mistake. It is an empirical hypothesis, perhaps an especially promising one, but still an hypothesis that may ultimately prove to be false. If it is rejected, then Harman's argument contra Davidson collapses. TCS for a language used in communication may be all we ever get or need by way of a theory of meaning. Since the language of thought hypothesis is empirical, the exact nature of the language or languages of thought remain to be specified by the development of cognitive psychology. This poses two possible difficulties for Harman's view embodied in his principles I and II. First, it may turn out that the language of thought postulated by cognitive psychology does not support I, that the meanings of linguistic expression are determined by the contents of the thoughts they are used to express. It could turn out that sentences in Mentalese underdetermine the meanings of public language. I will discuss this possibility

in Section 4. Second, it may turn out that in characterizing the language of thought cognitive psychology will employ as primitive a notion of linguistic meaning. I am not sure how this might happen but Davidson has given some arguments which seem to point in this direction (cf. [3]). He has argued that "... a creature must be a member of a speech community if it is to have the concept of a belief" and that "Someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error—true belief and false belief". If Davidson is correct then an account of belief and, if 3 is correct, of the language of thought will presuppose concepts of linguistic meaning. This is not the place to try to provide an assessment of Davidson's argument. I mention it only to make the point that Harman's views that language used in thought is more basic than language used in communication and that truth conditional semantics must be based on nontruth conditional semantics for the language of thought are not inevitable.

3 Harman conceives of CRS as providing an answer to the question: what makes something a concept with the content *C*? For example, he writes: "What makes something the concept red is in part the way the concept is involved in the perception of red objects in the external world" (p. 247). CRS's answer to this question is that an expression of a particular person *P*'s Mentalese is a concept with content *C* in virtue of playing a certain role in *P*'s psychology. There are two questions I want to consider about this view: 1. What is it to specify the conceptual role of an expression of *P*'s Mentalese? and 2. How are the conceptual roles of expressions of Mentalese related to reference and truth conditions?

CRS can be developed either atomistically or holistically. Procedural semantics, currently fashionable in AI, is an example of the former approach (cf. [18]). According to procedural semantics there is a stock of primitive expressions and various devices for constructing complex expressions. Associated with the former are procedures for determining whether or not an individual (presented in some canonical way) falls under the extension of the expression. Associated with the latter are algorithms for constructing complex procedures. Fodor [10] has pointed out affinities between procedural semantics and "old-fashioned verificationism" and argued that the well-known ills of the latter are inherited by the former. This does not seem to affect Harman's account since it is clear that he favors the holistic approach. However, he says very little about how the details of CRS are to be developed. The only detailed model of CRS which has been suggested is Hartry Field's probabilistic semantics (cf. [7]). But Harman rejects Field's account because it is unrealistic "since keeping track of probabilities involves memory and calculating capacities which are exponentially exploding functions of the number of logically unrelated propositions involved" (p. 247).

Perhaps Harman does not spell out the details of CRS because he correctly believes that CRS is an empirical hypothesis whose exact form awaits developments in cognitive psychology. However, I think that he does suppose that CRS will have the following general form. CRS will be a component of a functionalist theory of mind. A functionalist psychological theory

contains a specification of possible inputs (experience), a specification of possible outputs (behavior), a specification of mental states, and various lawful statements (perhaps probabilistic laws) linking them. According to the language of thought hypothesis, a specification of mental states, or a subset of the states, requires supposing that they have a syntax and semantics. Now the central claim of CRS is that the content of an expression of the language of thought has been specified when one specifies all of this theory. Harman's version is holistic since specifying the content of an expression involves spelling out its connections and nonconnections with all other expressions as well as with inputs and outputs.

Depiste the sketchiness of my (and Harman's) account of CRS we can see that it does fill certain lacunae left unfilled by TCS. We saw in Section 2 that, assuming the language of thought account of belief, Harman's argument shows that TCS cannot explain what it is to understand Mentalese. The question of how one understands the language one thinks in does seem to be a peculiar one. Fodor has remarked that it doesn't even make sense. CRS clarifies the situation. It is plausible that understanding a certain concept involves being able to use that concept appropriately. For example, to understand the concept red is, in part, to be able to discriminate red things. According to CRS an expression in *P*'s Mentalese has the content of the concept red just in case it plays the appropriate role in *P*'s psychology, including his discriminating red things. It follows that if some expression of *P*'s Mentalese is the concept red then *P* automatically understands it. The answer may appear to be a bit trivial—*P* understands the expression of his Mentalese since if he didn't it wouldn't be his Mentalese—but it is the correct answer. If there are any doubts compare the question we have been considering with "In virtue of what does a computer 'understand' the language it computes in?" Of course the understanding involved in understanding Mentalese is different from the understanding one has of a public language. I argued that understanding the latter involves knowing truth conditions. Not only would knowledge of truth conditions contribute nothing to explaining how we understand Mentalese but, it is clear, we do not know the truth conditions of Mentalese sentences. (Or, for that matter, even the syntax of Mentalese.) If *P* were to encounter a sentence of Mentalese written on the wall (in contrast to its being in just the right place in his brain) he wouldn't have the vaguest idea of what it means because he does not know its truth conditions.

There are certain sentences which seem to create difficulties for the truth conditional account of understanding when it is coupled with the language of thought account of belief. Suppose with Kripke that "Hesperus" and "Phosphorus" are rigid designators of Venus. Then "Hesperus is too hot to sustain life" and "Phosphorus is too hot to sustain life" have identical truth conditions. If there is a language of thought, then the truth conditions of these sentences are represented by some sentence *S* of Mentalese. It would seem to follow that anyone who understands both sentences and believes that "Hesperus is too hot to sustain life" is true will also believe, if he follows the consequences of his beliefs, that "Phosphorus is too hot to sustain life" is true. He should be as willing to assert one as the other. But this is certainly wrong. The obvious way out is to suppose that the two sentences express

beliefs which have different representations in Mentalese. But then there must be more to the semantics of Mentalese sentences than their truth conditions. According to CRS the two representations may differ in conceptual role even if they have the same truth conditions. It might be part of the conceptual role of the thought that Hesperus is too hot to sustain life to tend to produce, under certain conditions, utterances of "Hesperus is too hot to sustain life" but not utterances of "Phosphorus is too hot to sustain life". If this is correct, then even if one holds that truth conditional semantics provides the core of an account of language understanding it will need to be supplemented with an account of conceptual roles.

4 In this section I will discuss in a bit more detail the relations between CRS and TCS. Harman's view expressed in I and II is that the truth conditions of an indicative sentence are determined by its conceptual role. As I have already mentioned he also considers sentences which express truth conditions to be trivial and does not see a theory which specifies truth conditions as making much of a contribution to the theory of meaning. The issues involved in the relations between CRS and TCS come to the fore when one considers Putnam's twin earth stories (cf. [17]).

Putnam asks us to imagine a twin-earth which is just like earth except that what they call 'water' there is a liquid whose molecular structure is XYZ instead of H₂O. According to Putnam the reference of 'water' in English is H₂O while the reference of 'water' in Twin-english is XYZ. He uses this example to argue that meaning cannot be both what is in the head and what determines reference. Presumably the representation of H₂O in an earthling's head and the representation of XYZ in his twin-earth doppelganger's head are of the same type, since the two are physically type identical. Since the references of these representations differ they do not determine reference. If meaning does determine reference it is not in the head.

One way of reacting to Putnam's example is to deny that their 'water' and ours differ in reference. But Harman accepts this intuition of Putnam's. The example seems to create difficulties for CRS because on natural understandings of what is to count as conceptual role it would seem that the conceptual roles of the earthling's concept water and his doppelganger's concept water are the same but their contents are different, thus violating claim II.

So far as I can see there are two ways out of the twin-earth problem for an advocate of CRS who accepts Putnam's intuitions about reference. The first way is to include in the psychological theory that characterizes an organism's Mentalese microphysical descriptions of possible inputs. Then the content of a concept would be specified by specifying the theory and the inputs the organism actually experienced. The content of the concepts water on earth and on twin earth would be different since the respective concepts developed in the two organisms in response to different kinds of inputs, H₂O inputs on earth, XYZ inputs on twin-earth.

This response to the problem does not seem very attractive from the perspective of CRS. As I previously urged, the language of thought hypothesis is an empirical psychological hypothesis. If Fodor is correct in his claim

that psychological theory will be 'methodologically solipsistic' [9] then the theory developed by psychologists will not characterize inputs in terms of their microphysical structures. Methodological solipsism may turn out to be mistaken but it is difficult to see its rejection by CRS in response to the Putnam problem as anything other than an ad hoc maneuver to try to save CRS.

In any case Harman's response (p. 248) is a different one. He revises his position by replacing II by the principle that the content of a concept is determined by its functional role *relative to some normal context*. Even though the conceptual roles of the concept water is the same on twin-earth as on earth their contents differ since contents are evaluated relative to normal context and the normal contexts on earth and twin-earth differ. This revision is an appropriate one. In any case context needs to be taken into account to handle thoughts that we express using indexicals. However, this revision may be more far-reaching than Harman realizes.

Among the features of context which may need to be specified in order to fix the content of a thought are, at least, the thinker, the time of the thought, the references of demonstrative elements in the thought, and, if Putnam is correct, certain microphysical facts about the thinker's 'normal' environment. But it might turn out that context has an even heavier burden to bear. It will depend on how the psychological theory which characterizes the language of thought specifies inputs and outputs. If inputs are specified as, say, patterns of neural activity in the optic nerve, then it might very well be that in order to fix the content of a sentence of Mentalese we will need a great deal of information about the physical objects which caused particular patterns of neural activity in the optic nerve.

By revising II to accommodate Putnam's example, Harman is recognizing that Mentalese sentences have truth conditions. Given a complete psychological theory for *P* the conceptual roles of *P*'s Mentalese sentences are characterized. Conceptual role + context yields truth conditions. But from the perspective of CRS it is difficult to discern why we should be interested in truth conditions at all. Harman seems to think of truth conditions as epiphenomenal, the by-product of possessing the concept of truth and quotation but not themselves as making any contribution to semantics. But this distorts the relation between conceptual role and truth conditions. As long as we confine our interests to explaining the behavior of an organism for which we possess a psychological theory of the sort Harman hopes for, we will not need to employ the truth conditions of its Mentalese sentences or even the fact that they have truth conditions. But this does not show that truth conditions have no semantic function. I will give two reasons for believing that they are essential.

First, there are certain relations between a believing organism and its environment which involve truth conditions of the organism's beliefs. For example, if there is a fox in the vicinity then a rabbit in good working order is likely to come to believe that there is a fox in the vicinity. It is not unusual to explain why it is that someone believes that *q* by pointing out that *q* is the case and that he is appropriately situated with respect to the situation *q*. If *S* is the sentence of *P*'s Mentalese that he is related to when he believes that *q* then the truth conditions of *S*—viz. *S* is true iff *q*—can play a part in an

explanation of *P*'s acquisition of the belief that *q*. By neglecting truth conditions CRS misses an important feature of belief states, that they are information-bearing states.² Under suitable conditions *P*'s believing that *q* carries the information that *q*. This feature of belief is intimately related to the account of communication and understanding sketched in Section 1. Arabella's assertions about the weather are reliable because she asserts what she believes and her beliefs carry information about the weather. Truth conditions enter the picture because her belief that it's snowing carries the information that it's snowing in virtue of having the truth conditions that it's snowing. An account of why it is that Barbarella believes that it's snowing involves not only the truth conditions of Arabella's utterance "Es schneit" but also the truth conditions of her belief state.

Suppose that we have a psychological theory of the rabbit mentioned a paragraph back which is the sort of theory that Harman envisions as characterizing the rabbit's mental states. Assuming that the rabbit has beliefs and that the language of thought hypothesis holds for them, the theory will provide CRS for *Rabbitese*. We can use this theory to explain why the rabbit bears the belief relation to certain sentences of *Rabbitese*, for example the sentence we translate by "There is a fox in the vicinity". The explanation would cite the rabbit's mental states at some time previous to his acquiring the belief and certain inputs to his sensory system. Given this information the psychological theory will imply, or perhaps only make probable, that the rabbit acquires the belief that there is a fox in the vicinity. How does this explanation compare with the one that employs the truth conditions of the belief that there is a fox in the vicinity and cites the fact that there is a fox in the vicinity? On the positive side, it is much more specific and informative. The truth condition explanation is at best a rough one since rabbits don't invariably believe that foxes are present when they are. On the negative side, it is a much more complicated explanation. For most purposes the gain in precision probably will not offset the loss due to complexity. In any case the two explanations are not in competition. They are related to each other in something like the way in which a higher level functional psychological explanation is related to a lower level neurophysiological explanation. In our example the lower level psychological account helps to explain how the higher level account that appeals to truth conditions works. But even if we had a complete psychological theory of rabbits it would not render the higher level explanation useless or redundant. We would still need it to understand why the rabbit believes what he does.

My second reason for saying that truth conditional semantics is central to a semantical account of *Mentalese* is a development of the first reason. It is no accident that generalizations connecting situations in the world with an organism's beliefs via those beliefs' truth conditions hold. Rabbits and other believers are constructed so that when in their proper environments they acquire mostly true beliefs. Dennett has forcefully argued this point in a number of recent papers³ sometimes emphasizing a conceptual necessity and sometimes emphasizing a sort of evolutionary necessity between being a believer and believing truths. In either case Dennett's claim is that "true believers mainly believe truths".⁴ If this is correct then the fact that beliefs,

or sentences of Mentalese, have truth conditions is central to our notion of belief.

I am not sure that Harman would disagree very much with this. He never denies that Mentalese sentences have truth conditions. But in the course of his advocacy of CRS he does claim that the possession of truth conditions is a rather trivial matter. In my comments I have endeavored to show that, on the contrary, truth conditions are at the very heart of semantical theory.

NOTES

1. Strictly speaking the required knowledge is that "Es schneit" is true when uttered at time t in location w iff it's snowing at t in the vicinity of w .
2. Dretske in [5] develops an information-based account of belief. I criticize his account in a review forthcoming in *Philosophy of Science*.
3. See Dennett [4] and "True believers", a possibly penultimate draft in circulation.
4. From "True believers", abovementioned draft, p. 8.

REFERENCES

- [1] Davidson, D., "Truth and meaning," *Synthese*, vol. 17 (1967), pp. 304-323.
- [2] Davidson, D., "Belief and the basis of meaning," *Synthese*, vol. 27 (1974), pp. 309-323.
- [3] Davidson, D., "Thought and talk," in *Mind and Language*, ed., S. Guttenplan, Oxford University Press, Oxford, 1975.
- [4] Dennett, D., "Three kinds of intentional psychology," in *Reduction, Time, and Reality*, ed., Healy, Cambridge University Press, Cambridge, 1981.
- [5] Dretske, F., *Knowledge and the Flow of Information*, M.I.T. Press, Cambridge, Massachusetts, 1980.
- [6] Dummett, M., "What is a theory of meaning, part one," in *Mind and Language*, ed., S. Guttenplan, Oxford University Press, Oxford, 1975.
- [7] Field, H., "Logic, meaning, and conceptual role," *Journal of Philosophy*, vol. 74 (1977), pp. 379-408.
- [8] Field, H., "Mental representations," *Erkenntnis*, vol. 13 (1978), pp. 9-61.
- [9] Fodor, J., "Methodological solipsism," in *Representations*, M.I.T. Press, Cambridge, Massachusetts, 1981.
- [10] Fodor, J., "Tom Swift and his procedural grandmother," in *Representations*, M.I.T. Press, Cambridge, Massachusetts, 1981.
- [11] Harman, G., "Meaning and semantics," in *Semantics and Philosophy*, eds., M. Munitz and P. Unger, New York University Press, New York, 1972.
- [12] Harman, G., "Language, thought, and communication," in *Language, Mind, and Knowledge*, ed., K. Gunderson, University of Minnesota Press, Minneapolis, Minnesota, 1975.

- [13] Harman, G., *Thought*, Princeton University Press, Princeton, New Jersey, 1975.
- [14] Harman, G., "Conceptual role semantics," *Notre Dame Journal of Formal Logic*, vol. 23 (1982), pp. 242-257.
- [15] Lepore, E. and B. Loewer, "Translational semantics," *Synthese*, vol. 48 (1981), pp. 121-133.
- [16] Lewis, D., "General semantics," *Synthese*, vol. 27 (1974), pp. 18-67.
- [17] Putnam, H., "The meaning of meaning," in *Language and Reality* (Collected Paper, vol. 2), Cambridge University Press, Cambridge, 1978.
- [18] Suppes, P., "Procedural semantics," in *Proceedings of the 1980 Wittgenstein Conference*, forthcoming.

Department of Philosophy
University of South Carolina
Columbia, South Carolina 29208