

KNOWLEDGE AND UNDERSTANDING

BY JOHN CAMPBELL

I

Is linguistic understanding a form of knowledge? Should a theory of understanding specify the content of *knowledge* constitutive of language-mastery? An affirmative reply to these questions obviously constitutes an attractive basis for semantics and linguistic theory. My aim in this paper is to supply a foundation for that basis, in the face of recent attempts to undermine it.

In comprehending speech-acts couched in one's mother tongue, there is typically no need for conscious deliberation over their construal. One simply perceives them as having the meaning and grammatical structure that they do. In itself this raises no difficulty for the view that understanding is cognitive. Just as I may perceive a piece of music differently when I know its place in musical history, so linguistic perception may be viewed as a result of mastery's being a kind of knowledge. It has, however, recently been given central place by cognitivism's opponents.

According to John McDowell, possession of such a perceptual capacity is constitutive of mastery of a language. The aim of a theory of understanding, on this view, is to specify, for every possible speech-act of the language, the meaning which one who understands the language will thereby perceive it as having. And the theory must not attempt to explain that capacity in terms of underlying knowledge ([9], 165-9). However, McDowell does not say whether, on his account, possession of the capacity is wholly, or partially, constitutive of understanding. And I believe that problems arise no matter which line is taken. Mastery of a language typically enables one not only to comprehend its speech-acts, but also to perform them. Any view which maintains that perfectly normal individuals may have mastery of a language yet be simply incapable (and not because of any deficiency in their physical productive powers) of free unreflective speech in that language, is surely not credible. And so, it would appear that the perceptual capacity on which McDowell concentrates cannot plausibly be regarded as *wholly* constitutive of speaker competence.

What, then, would a completed McDowellian theory of understanding look like? It would have to characterise at least two capacities: to *comprehend* perceived speech-acts, and to *perform* them. And it seems that these capacities would be thought of as intrinsically unrelated, with no explanatory structure typically underlying possession of both.

Yet we simply do not find people who, though not deficient in rationality or perception, are quite able to express their thoughts in their native tongue

but are simply incapable of comprehending speech-acts couched in that language; or vice versa. Moreover, it seems that, by and large, the operation of some *one* process is all that is required for people to acquire these two capacities. The reason seems obvious. Some one thing — some one structural property of a person — typically grounds his possession of *either* capacity. For McDowell, however, the facts for which this explanation accounts seem destined to remain opaque.

There appears, then, to be good reason to try to delineate a structure constitutive of understanding, which would reveal both capacities to be *consequences* of language-mastery. But why should we think that the structure in question has to be described specifically in terms of *knowledge*? The important point seems to be that those who understand a language are thereby capable of learning, on occasion, by dint of their own sensory experience and reasoning, that its sentences are true; they are able to know their truth. If the intensional basis of one's interpretation of a sentence were epistemically unsound, if it comprised merely wishful thinking, or accidentally true beliefs, or indeed anything other than knowledge as to how the language is to be understood, then it would be simply incomprehensible how, on such a basis, one could *know* the truth of a sentence, or its falsity; we could not explain our own epistemic abilities.

Once we admit the legitimacy of seeking a unified explanation of the intensional phenomenon of the concomitance of speech-productive and perceptual linguistic capacities, we are driven to look for an intensional explanation of them; for neural explanations seem to serve only to account for physically described effects. And then, if we are not immediately to be forced to radical scepticism about knowledge of the truth of sentences, we appear to be compelled to characterise language-mastery as a form of knowledge: for no other intensional state seems capable of supplying a sufficiently firm epistemic base for our manifest cognitive powers here.

Furthermore, comprehension of our language serves as an epistemic basis for the acquisition of knowledge of other types. Those with a certain minimum of rational creative ingenuity are capable of exploiting their language-mastery to develop an open-ended variety of sophisticated indirect tests for finding the truth-values of various sentences. The same points may also clearly be made about knowledge acquired through hearsay or testimony. Now of course, it may emerge that there just is no explanatory structure of the type which we seek; that McDowell is right and that all there is in common between most of those who have mastered a language is the perceptual capacity he describes. Our conclusion is only that there is, at the moment, reason to look for a deeper account.

II

According to Michael Dummett, a theory of understanding should proceed by characterising that knowledge whose possession is constitutive of

mastery of the language in question. Moreover, understanding is held to be a behavioural matter. Dummett connects these theses by this demand: for every statement (in a theory of meaning) which specifies the interpretation of a sentence, a theory of understanding must provide a behavioural account of what it is to have implicit knowledge of that statement ([4], 67-76; cf. [5], 134). It may be wondered, however, whether we could not accept both that understanding is cognitive, and that mastery is behavioural, without imposing this reductionist requirement. But the upshot of this would be that the theory, while purporting to describe a form of behaviour, is in fact employing concepts which do not effect any systematic classification of behaviour at all. Such a theory would appear to be an exercise in futility. (Dummett suggests that a theory of this type would be holistic, and the above is the core of his objection to it. Cf. [3], 116; [4], 70-2.) But if it must be reductionist, the theory faces a formidable task; and it is not clear that it can be performed.

The theory has to distil an account of what understanding *alone* is. So it has to filter out all except those aspects of behaviour which manifest only a person's grasp of the language. Yet surely, it is only in conjunction with the causal operation of propositional attitudes that mastery manifests itself behaviourally. To take a simple example, an assent to a question may depend not only on the subject's comprehending the query, but on his desires to be sincere and forthcoming, and his beliefs about the world. So at the level of behaviour there appears, at first sight at any rate, to be nothing salient for the theory to distil. Every aspect of behaviour which looks as if it might be relevant is coloured by the subject's beliefs and desires, which are extraneous to his understanding.

In the face of this fact, there are two general ways in which one might try to supply a behaviourist description of understanding. But I cannot see how to make either of them work. One would be to give a general formula which yields, for any set of propositional attitudes, a behavioural description of what it is for someone with those attitudes to have a grasp of the language. Now, the only way I can see of doing this is by redefining the role of the concept of knowledge in the account. Thus, the theory will characterise knowledge whose possession is not constitutive of, but merely sufficient for, mastery. Possession of that knowledge will be seen not as a behavioural matter, but as a causal factor underlying behaviour. And the formula will be this: someone with a particular set of propositional attitudes understands the language just in case his behaviour is the same as that of someone with those attitudes *plus* the knowledge characterised by the theory.

It is clearly crucial to this approach that the knowledge in question is not seen as part of the common property of masters of the language. For the theory aims to sustain behaviourism; but possession of that knowledge must be seen as a causal factor underlying behaviour, and not itself a behavioural matter. It is after all only in virtue of its causal role that the

knowledge serves to provide an adequately relativised description of behaviour-patterns. So typically, although a master of the language will behave *as if* he has that knowledge, it will not be possession of it which explains his behaviour, but merely some internal neural structure. The problem here becomes apparent, however, when we consider that propositional attitudes ordinarily vary over time. For the neural structure then turns out to have a certain amazing property: it *interlocks* and *combines* with the agent's beliefs and desires to produce his behaviour *exactly* as would a certain propositional-attitude complex: namely, the knowledge specified by the theory of understanding. There is surely but one explanation of this phenomenon: but the abandonment of behaviourism appears to be the price of accepting it.

The other approach to providing a behaviourist description of understanding is to employ the way of brute force. The theory will supply an exhaustive statement of the sets of propositional attitudes a subject may hold, perhaps allowing also for variations in subjective probability and utility assignments. And for each such set, it will characterise behaviourally what it is for someone with that combination of propositional attitudes to understand the language. However, while such an account purports to tell us what understanding is, it can offer no explanation of the fact that mastery of a language is something which once acquired is typically *retained* through alterations in the subject's beliefs and desires; for it postulates no continuous underlying structure constitutive of understanding. One might reply that fortunately, those who at some point acquire understanding do also acquire a neural structure which not only interacts with their current propositional attitudes to yield the appropriate behaviour-pattern, but can also be relied upon to interact appropriately with updated stocks of beliefs and desires as they develop over time. This neural structure, too, however, clearly has to be capable of interlocking and combining with a subject's beliefs and desires to produce his linguistic behaviour exactly as would a propositional-attitude complex. And in that case we would do better to forget the behaviour and concentrate on characterising the complex. There is, so far as I can see, no way round the difficulties for the project of supplying a behaviourist account of mastery; I propose, therefore, that we consider the prospects for a cognitivism which employs a non-behaviourist conception of knowledge.

III

Many philosophers have insisted that a description of understanding must be compositional; that it must display the meanings of sentences as systematically dependent upon the meanings of their parts. And the philosophical interest of a description of language-mastery is to a great extent dependent upon its being, in this sense, structure-discerning. For languages do not have semantic features which are somehow independent of the way in which they are understood by their speakers. Now, one aim of (non-behaviourist) cognitivism is to expand the range of propositional-attitude

explanation; to enlarge (*inter alia*) the sphere of recognisably intentional (which is not to say, reflective) action. For it is by giving such an explanation of an action that we show it to be intentional. And the aim is to do this for various aspects of linguistic behaviour, such as the speaker's choice of words in speech, or his reactions to heard utterances. Yet at the level of action-explanations, we shall often be concerned with the speaker's knowledge of the meaning of this or that *sentence*. So one way for the theory to proceed would be by specifying the content of knowledge held to constitute understanding of each sentence of the language. Some urgency thus attaches to the question raised by Crispin Wright [12]: under what circumstances should an account of language-mastery be compositional? What would force compositionality upon us?

We have seen, however, that one aim of a cognitivist description of understanding ought to be to explain the capacities of speakers to perceive the meanings of heard sentences. Now, it seems plausible to suppose that we perceive heard sentences as fragmented into words in a certain way; and this seems essential for unreflective comprehension of them. Think, for example, of the familiar experience of hearing someone talking heavily accented English which is simply unintelligible until one starts to pick out the clauses and words, to hear the structure in what he says. To *perceive* a speech-act as having a certain meaning, one must surely perceive it as having a certain structure; as being organised out of constituent words in some one way. Such a hypothesis clearly favours a compositional account of mastery. For if understanding were explained solely on a sentence-by-sentence basis, then we could not say why perception of structure should be at all relevant to perception of meaning.

We find a rich and satisfying explanation of the phenomenon, however, if we suppose that the knowledge constitutive of understanding relates primarily to words and their composition into sentences. For then, it will not be enough for comprehension of a heard sentence that one be master of the language in which it is couched. For, it is not enough that one know what sentence one has heard, and know the significance of its words, and the significance of the manner in which they are in fact organised in the sentence. For comprehension, one will need also to know that that heard sentence *is* compounded from those words in just that manner. And that latter knowledge is exactly what one unreflectively obtains only by trusting one's perception of the structure of the sentence. We sought an explanation of why perception of meaning should depend upon perception of structure, and now we have one. I shall argue that in fact, such an explanation can only be supplied by a knowledge-based approach to understanding.

There is further motivation for compositionality in a cognitive account of understanding. Thus consider that one central practical test for the intelligent use of language is the ability to explain the meanings of the words used. This is the criterion we employ whenever — as for example in

examinations — we want to know whether someone is merely repeating heard noises. Of course some people are more articulate than others; we are not always capable of delivering polished paraphrases of our words. Whether they use ostension, metaphor or paraphrase, however, the general run of speakers are quite capable of supplying explanations of the meanings of words which they employ. The point is not confined to terms learnt by verbal definition: there is, for example, a good deal to be said about the meanings of colour terms, drawing out the intricate relations between them, and any competent speaker can at least begin to articulate some of the relevant information here. And surely only an account of mastery which takes it to consist in knowledge of the meanings of terms, rather than as relating directly to sentences, can concede the possibility of this phenomenon.

We should remark also that those who understand a language are often capable of comprehending words outwith the context of sentences; for we often use individual words as *elliptical* for full sentences. This imposes on our audience the burden of recognising *which* sentence the solitary word is elliptical for. Suppose that, as the nervous guest stands petrified by the appearance of a snake at a party, the host murmurs 'Stuffed'. One's grasp of the meaning of the individual word used is *immediate*: we recognise immediately that it is being said of something that it is stuffed and, in the total context, look around for something that may reasonably be being said to be stuffed, alighting on the snake. The moral is that our language-mastery does not relate purely to grasping the meanings of sentences. For even if speech-acts involving solitary words must invariably be elliptical for sentential speech-acts, still a direct grasp of the meaning of the solitary word may be an essential datum in realising what the relevant sentence is. The point is, I think, too easily obscured by reliance on the dictum that it is only in the context of a sentence that words have any meaning.

Gareth Evans [6] supplies further motivation for compositionality. If one who has had explained to him certain sentences containing a new word is thereby capable of comprehending other (new) sentences containing the word, then his understanding must surely be word-oriented. Or again, if someone who *loses* competence with certain sentences containing some word loses competence with *all* sentences containing that word, then a compositional account offers the obvious explanation of the situation. (As we shall see, Evans does not give a cognitivist reading of the upshot of these remarks; yet plainly the cognitivist may take them to show a need for discernment of structure.)

It seems, then, that compositionality has a powerful explanatory role to play within the context of a knowledge-based approach to understanding. It is an open question whether any point can be assigned, on any other approach, to the idea of displaying the composition of a sentence (cf., Wright [12]). It seems also to be an open question whether any alternative to cognitivism can explain the data for which compositionality accounts within

the context of a knowledge-based approach. To bring out some of the difficulties here, I shall deal with two recent accounts of structure-discernment.

The natural way to state the knowledge described by a cognitivist view of understanding is by an axiom-system: the axioms specifying the meanings of words of the language, the theorems (*inter alia*) specifying the meanings of sentences. Now many who shun cognitivism would favour the axiomatic approach; but it is really not clear why. According to McDowell [10], the axioms are merely convenient computational devices for entailing the theorems; the account makes contact with reality, in particular with language-mastery, not at the level of the axioms, but of the theorems. McDowell insists on the axioms; it is an *a priori* constraint on his account that it discern structure. But why? All the theory has to do, on McDowell's view, is to specify the perceptual capacity he takes to be (at least partly) constitutive of competence. And, complications about indexicals and force aside, it will have done this, for indicative sentences at least, when it delivers the result that, for every sentence, ' ϕ ', competent speakers perceive utterances of ' ϕ ' as assertions that ϕ . So on the face of it, it would seem that a great deal of pointless labour would be saved if, in specifying his perceptual capacity, McDowell contented himself with the infinitary schema: ' ϕ ' is T iff ϕ . The demand for axiom/theorem presentation seems, on this approach, to have nothing whatsoever to do with the description of mastery. It might be replied that McDowell wishes to take into account the fact that we typically perceive heard sentences as having a certain structure. Yet it is not at all clear *why* he should want to do so. For there appears on this account to be no reason whatsoever why perception of structure should be essential for perception of meaning. And if perception of structure is really a peripheral epiphenomenon in our use of language, then it is totally gratuitous to demand that a description of understanding should deal with it. Thus we have yet to see what objection there could be to the use of the infinitary schema, on this approach to language-mastery.

Suppose that the demand that a description of mastery characterise the structures perceived by speakers were, for whatever reason, legitimate. Even this would not in itself yield the result which McDowell desires. For there is as yet no motivation for insistence on the *derivability* from axioms of theorems which deal with the meanings of sentences. For we have no reason yet to think that the structure in question is *semantic*. To see this, remember that for McDowell there appears to be no reason whatsoever why there should be any essential connection between perception of structure and perception of meaning. His position, therefore, seems to offer no objection to an account which describes the structures speakers typically perceive in sentences quite independently of a description of perception of meaning. For this latter task, we might employ the infinitary schema. On this view,

any connection between the two descriptions would be purely fortuitous. Of course, this view is not McDowell's; but it is not obvious that he can say why not without appealing to a compositional knowledge-based conception of understanding.

In reply to Wright's challenge to motivate compositionality, Evans [6] makes the following suggestion: the discernment of structure by a description of understanding is in effect a discernment of non-psychic, purely *neural* structure. The leading idea can be put crudely as follows. Suppose the theory identifies *w* as a word. That identification is correct just in case for any speaker of the language, there is a partial neural cause common to all of his perceptions of the meanings of heard sentences containing *w*. Thus the total fragmentation of the language into words and their modes of organisation into sentences constitutes a complex hypothesis about the structure of a network of states which are partial causes of sentence-comprehension.

Now the resultant view of semantic structure may seem to provide a vindication of McDowell's views; it fits in beautifully with his opposition to the insistence that cognitive structure be discerned. For it seems to offer structure without making any essential use of the notion of knowledge. It is true that Evans employs the notion of "tacit knowledge" in stating his view, but its use is dispensable. The theory could be stated as above and elaborated as a purely neural hypothesis which makes no essential use of any psychological concepts other than that of comprehension, or perception of the meaning of, a sentence. Indeed, it may serve to explain the word-by-word acquisition of language, and the point that loss of linguistic competence may be expected to be a word-by-word matter too. For on this account, the neural structure which realises language-mastery is composed of a battery of states, each of which is a partial cause of the comprehension of a wide range of sentences. So the instillation of a new state of this type will, in co-operation with other, already present states, typically result in a capacity for comprehension of a wide range of new sentences. Destruction of any one such state will, since all the rest are merely partial causes of sentence-comprehension, result in loss of competence with all those sentences in the causation of whose comprehension that state was a necessary factor.

We should, however, have considerable reservations about this apparently attractive view. For it denies much that we should expect the compositionality of a description of understanding to explain. Can it explain why perception of structure should be essential for perception of meaning? Evans insists that his theory can give an explanation here; but in this he is surely wrong. The key point is that on this view, episodes of perception of meaning have no explanation, though they are caused. That sounds paradoxical, but in fact the point is familiar.

There is, on this view, no *psychological* machinery which typically explains a speaker's perception of the meaning of a heard sentence. Now a perception

of meaning may well be identical with a physical event. And that physical event doubtless has a neural explanation; that is what is meant by saying that the perception has a physical cause. It is, however, fallacious to infer from this that *qua* perception of meaning, it has an explanation. For the neural event might explain the right event, but under the wrong description. As Evans stresses, and this is exactly what opposition to cognitivism amounts to, the perception of meaning is held to have no explanation in psychological terms. But surely *qua* perception it could only be given an explanation in those terms. It thus transpires that on Evans' approach, perceptions of meaning have no psychological explanation whatsoever.

So perception of structure has no role to play either in the explanation of perceptions of meaning physically described — for here it is the purely neural account that is relevant — or in the explanation of such episodes described as perceptions — for in that guise they have no explanation at all. Perception of structure thus turns out, on this view, to be essentially irrelevant to perception of meaning; it is at best an epiphenomenon in the use of language. And so with Evans we are no closer than we were with McDowell to an explanation of why perception of structure is essential to perception of meaning. As noted earlier, it is true that Evans introduces his theory by using the words 'tacit knowledge'; but this cannot help here. For all that distinguishes Evans' view from a knowledge-based explanation of understanding is his insistence that such talk is merely a convenient *façon de parler* in discussing neural states; it is to be taken with a pinch of salt.

There are, indeed, further problems here. Evans' account cannot explain — and, it seems, must rather deny — two phenomena alluded to earlier: our capacity to explain the meanings of the individual words we use, and our capacity to recognise for which sentence an utterance of a solitary word is elliptical. Again, it is natural to suspect that justice to the phenomena can only be done by a knowledge-based explanation of mastery.

There is one type of response which may save an Evans-type account here; but it is important to beware of its pitfalls. The idea would be to enrich the characterisation of the components of the neural structure realising mastery. Evans was content to assign to these states the causal role alone of producing comprehension of sentences. But we might try to go further, and insist that these states invariably also either realise or cause perception of structure in producing perception of meaning, in the hope of explaining away the fact that the former appears to be essential for the latter. We might insist also that these states produce in the subject the capacity to explain the meanings of the words he uses, and the ability to understand solitary words used as elliptical for sentences. The problem with such *ad lib* enrichment is that the neural states described grow progressively harder to distinguish from intensional states. It becomes more and more difficult to see how they could have the properties ascribed to them unless they were intensional. And so we may end by being committed to cognitivism.

To this Evans points out that linguistic knowledge is ordinarily of use to its possessor only in speaking and understanding. This merely displays, however, the oblique relation of the content of such knowledge to many characteristic human interests: knowledge of the plot structure of *Bleak House*, for example, is still knowledge though its use is limited.

Again, Evans holds that linguistic knowledge cannot combine with the subject's thought in general to yield new beliefs. Yet the pervasiveness of context dependence means that linguistic knowledge is often not sufficient for speech comprehension. Understanding of a speech-act is typically derived jointly from linguistic and non-linguistic knowledge. As Bar-Hillel stressed, there is no demarcating the salient information in advance. Something as unexpected as knowledge of the sizes and shapes of play pens and writing pens may be needed to construe 'The box was in the pen'. The alleged inferential insulation of semantic knowledge is, then, a myth.

IV

It thus appears that a knowledge-based approach offers our best hope of achieving insight into language-mastery. What are the problems for such an approach? Potentially the most decisive criticism, it seems to me, is that such an explanation must inevitably be circular. So we now ask where, exactly, the circularity is held to enter. It seems that the very same issues will arise if we view cognitivism as aiming to explain understanding as a type of belief. By doing so we may hope to avoid irrelevant complications about the evidential ancestry and truth of what one knows; so in this section I shall take that view. We cannot begin better than by distinguishing very roughly between three possible responses to the question: what is it to believe that ϕ ?

One view proceeds by drawing a sharp distinction between sentential and other beliefs. For present purposes a sentential belief is of the form: the belief that ' ϕ ' is true. The view is that belief generally is to be explained in terms of sentential belief. With this view we may contrast two others. First, it might be held that to believe that ϕ is to possess a neural state having a certain causal role. Functionalism might be seen as one version of this. Secondly, it might be held that it is simply wrong to seek any type of reduction or explanation of what it is to believe that ϕ . Now some important distinctions are being fudged here; but my aim is not to present three views which are necessarily exclusive; rather, they are to be *independent*, in that one could hold any one without subscribing to the others.

The thought behind the charge that cognitivism is circular must surely be that belief (or knowledge) is ultimately to be explained in terms of understanding. So the objection presupposes that a substantive answer can be given to the question: 'What is it to believe that ϕ ?', an answer which ascribes understanding of something (linguistic) to the believer. But understanding of what, precisely? The only remotely plausible candidate as a

reply is surely: understanding of 'φ' or some synonymous sentence. No other linguistic expression seems to be relevant. Yet clearly, such understanding is not sufficient for the belief that φ. And it would seem that the only plausible condition we can add to ensure sufficiency, while retaining the necessity of the clause that the believer must understand 'φ' or some synonym, is this: the subject must believe that 'φ' (or the synonym) is true. It thus transpires that the criticism of cognitivism as circular depends on the view that belief is to be explained in terms of sentential belief; that is, the first of the theses stated above.

Michael Dummett [3] has presented a version of the "circularity" charge which may at first sight seem to be free from such theoretical commitments. The cognitivist, Dummett points out, must distinguish between knowledge that the sentences of his semantic theory are true, and knowledge of what the theory states; that is, between knowing merely that 'φ' is true, and knowing that φ. Yet how is this distinction to be explained, except by noting that in the latter case, but not in the former, the speaker must understand 'φ' or some synonym? ([3], 110 ff.) At first, then, this argument appears not to depend on any contentious doctrine about the nature of belief or knowledge, but only on the demand that a distinction be explained. But what kind of explanation is required? Surely one could *communicate* to someone the gist of the distinction by, for example, pointing to the difference in explanatory role of belief and sentential belief. Thus a hungry monolingual Frenchman who knows only that 'Supper is on the table' is true is in a very different position from one who knows (by looking) that supper is on the table. Once we explain that one is, while the other is not, in a position to sate his hunger, surely anyone would catch on? Yet here there is nothing which only one of the two grasps. Dummett's reason for insisting that an explanation of the distinction must appeal to the notion of language-mastery must, so far as I can see, be that he requires a *deep* explanation of the distinction, rather than a cursory guide to its application. And a deep explanation will be one which tells us what it is to believe that φ, tells us what it is to believe that 'φ' is true and thereby makes manifest the difference between them. The thought is then that in giving such a deep explanation we shall have to employ the notion of understanding. But it is hard to see anything inevitable about this. It is certainly not obviously wrong to suppose that one might give an explanation, in terms of causal role, of what it is to believe that φ and of what it is to believe that 'φ' is true, without using 'understands'. The cognitivist who gave such an explanation would have shown his account to be free of circularity. Or again, it may be correct to insist that *no* deep explanation can be given of the distinction to which Dummett alludes. And then again there is no circularity in cognitivism. It is only when we assume that in explaining what it is to believe that φ, as part of giving a deep explanation of that distinction, one must ultimately appeal to the idea that the subject understands 'φ' or some synonym, that we see

where the circularity in cognitivism is supposed to enter. But now we are back with the idea that belief is to be explained in terms of sentential belief. Of course, none of this denies something to which Dummett is obviously appealing: namely, that to acquire belief or knowledge by testimony, one must understand the expressions used. We need not deny either that a necessary condition of believing that ϕ is understanding ' ϕ ' or some synonym. What is at issue is rather the question of what role, if any, these conditions are to play in *explaining* what it is to believe that ϕ .

It thus seems plausible that all versions of the criticism of cognitivism as circular or regressive depend on supposing that belief is to be explained in terms of sentential belief. Those who view the proponent of a knowledge-based approach to comprehension as committed to an infinite chain of homunculi within the speaker (the homunculi being those who understand the semantic sentences and believe them to be true), and those who view him as explaining understanding of one language in terms of understanding another, obviously fall into this camp; together, of course, with those who do not envisage the cognitivist as making these regressive moves but simply take his account to be circular. Let us, then, assess this worry. I shall not here essay a definite answer; the objection, however, seems to me to be at the moment less than decisive.

The most cautious reply to the criticism would be that it sets too high standards for a description of understanding. A cognitivist description of language-mastery, even if it were circular, would not leave us exactly where we began. Rather, we would have fragmented mastery into comprehension of individual words and their modes of composition, and *a fortiori*, analysed the structure of the language to some degree. What we would have done also would have been to *relate* understanding to other psychological states, rather than in some strict sense explained the former in terms of the latter. It should be obvious that as a response to the charge of circularity this has much polemical strength; for anyone who presses that criticism appears to be committed to the enterprise of explaining belief in terms of sentential belief. And that seems doomed to either circularity or regression.

Nevertheless, I suspect that we ought to pursue the path of considering whether language-mastery can be given a genuinely reductive explanation in terms of knowledge; such an explanation would clearly be highly satisfying. We could then both articulate in detail how understanding relates to the propositional attitudes in, e.g., action-explanation, and we could see the explanatory role of various components of understanding. We might thus gain greater insight into the structure not only of our mental states but of our language.

It is in any case not obvious that the theoretical presuppositions of the "circularity" criticism are correct. We should certainly feel sceptical about the play that is made with the notion of synonymy. Such play is essential, of course, if we are to acknowledge that monolingual speakers of different

languages may share beliefs. Yet in using the notion of synonymy to explain what it is to believe that ϕ , we apparently commit ourselves to employing it to explain when two propositional-attitude states are the same. And now it would seem that any explanation of synonymy must eschew the intensional. So we seem likely to end by giving behavioural or neural criteria of identity for beliefs, if we are not willing to take the notion of synonymy as primitive. We have yet to see a plausible set of such criteria, however.

Moreover, it is not clear that the proposed explanation of belief in terms of sentential belief is even extensionally correct. It is not clear that believing that ϕ is not a characteristic *consequence* of understanding ' ϕ ' or some synonym, and believing that it is true; rather than having to be explained as actually consisting of such understanding and belief. The point is that one has to *connect* the understanding and sentential belief and — as it were — draw a conclusion from them if one is to believe that ϕ . Thus, consider someone who hears catastrophically bad news — say, a mother who hears that her son has almost certainly been killed. Perhaps the statement derives from an authoritative source, so that whenever she considers the sentence 'N.N. has been reported missing in action', she is ready to admit its truth; it is just that whenever she does so, her mind goes blank as to its meaning. But when she considers the sentence without considering its truth, she is perfectly capable of comprehending it. It would be wrong to ascribe to her the belief that N.N. has been reported missing in action, if that belief is the basis of none of her actions. But if this description of the case is correct, then understanding and sentential belief are typically causes of the belief that ϕ , rather than being constitutive of it.

V

We have, I submit, found clear motivation for a knowledge-based approach to speaker competence; whether or not such an approach is in the end correct, it is surely worth serious investigation. I close now with a series of problems, theses, analogies and conjectures.

Notice, incidentally, that our arguments in §§ I-III were based on recognition of the complex way in which competence interacts with the web of the speaker's intensional states. It is perhaps worth pointing out that argument in this manner is not readily susceptible to *reductio* by analogy: it cannot be applied to show equally that homing pigeons have knowledge of trigonometry, for example; for the *initial* web of states which the argument needs is there absent or at best rudimentary.

(1) A knowledge-based approach to understanding may seem to require that we ascribe to subjects knowledge stateable only by using concepts which they do not possess; and secondly, that we ascribe to them knowledge of which they cannot give verbal expression. The central reason for current reservations about cognitivism is surely the apparent complexity of attempts

to give precise statements of knowledge which might be held to be constitutive of competence; for it often seems that whenever we manage to state knowledge which looks as though it would suffice for mastery, it turns out not to be necessary for competence. In dealing with the resultant twin uncertainties about conceptual complexity and ready verbalisability — whose distinctness will emerge more clearly — what is required, it seems to me, is careful attention to what, exactly, a cognitivist view of comprehension claims.

We do well to recognise immediately that not all knowledge need be readily verbalisable by those who possess it. Consider the case of knowledge of a geographical region. Suppose, for example, that Sam has detailed knowledge of the layout of Colditz. We may suppose that this is not merely a quasi-behavioural matter of his being able to find his way to and from selected spots using landmarks. Rather, he has a clear mental picture of the internal layout and external relations to each other of the various buildings. So he can, for example, devise new shortcuts and escape routes. Yet there is no reason to suppose that he is capable of giving ready verbal expression to this knowledge. Many of the shapes involved, for example, may have no names in English.

Knowledge of the melody of “Lilliburlero” supplies a case which is in some ways yet more striking. Such knowledge need not be merely a quasi-behavioural matter of being able to sing the tune or recognise it. One may be able to run the tune over in one’s head, devote some silent contemplation to the categorisation of its musical style, and compare it mentally to other songs and say which one prefers. Clearly, however, one may not only be incapable of giving a precise verbal statement of the content of one’s knowledge, but may also have difficulty in giving it any straightforward expression at all which would be readily accessible to an audience. The point is just that knowledge may not be manifestable by simple verbal formulation of it by the subject.

(2) There is constant pressure on the scientist to give precise verbal articulation of his hypotheses, with a view to communal assessment, elaboration and quantitative testing. Recent philosophy of science, however, has stressed the importance in scientific thought of models, metaphors and analogies (cf. Suppe [11] for a review discussion). From the wide variety of current uses of the term ‘model’, let us select one which seems to be of especial relevance: the idea of a model as a quasi-pictorial *image* of the structure of some phenomenon or process. For example, it seems fairly clear that in his more speculative discussions of the nature of light, Newton was working with an intuitive image of light as something both particulate and wavelike, a conception which he had difficulty articulating. A full and satisfactory account of scientific thought is only possible when we deal with both of its aspects, the verbally stated hypotheses and the intuitive models, and the interplay between them: the hypotheses being checked by appeal to

such models (cf. the controversies over "action at a distance"), and the models being progressively enriched and articulated into verbal hypotheses.

Now the topographical and musical examples given above seem to me to show conclusively that knowledge need not be readily verbalisable by those who possess it. Yet they make the point in what is for our purposes a rather unconstructive way; for knowing a language may not seem, at first glance, to be much like knowing an area or a song. But the analogy between possession of an intuitive model of some phenomenon and language-mastery seems immediately suggestive. And of course, the idea that there are significant parallels between scientific knowledge and linguistic understanding is hardly new.

Indeed, for those who work directly on empirical studies of language-learning, the parallel often seems inescapable; children do seem to proceed by forming *conjectures* about grammar and meaning. In applying the parallel, however, a suspect philosophy of science seems sometimes to be assumed. Thus, consider the following simplified version of Chomsky's argument for the existence of an innate "universal grammar". "Induction", as Chomsky uses the word, consists in the verbal formulation of hypotheses; in this case, about grammar and meaning, and doubtless unconsciously, in an unspoken language of thought (cf. Harman [7]). The hypotheses are of the definite, precise type found in formal grammars. The question is: what data are available in the construction of such hypotheses? Experiential data alone are simply not sufficient for this; therefore there must be an innate grammar which guides conjecturing (cf., e.g., [1], 40 ff.). Philosophers (see, e.g., Wright [12]) who feel wary of the notion of an unconscious verbal hypothesis are inclined simply to dismiss the analogy between the growth of scientific knowledge and language-learning at this point; a move which seems to me disastrous, because it leaves us with no idea of how we might construct a psychological account of language-acquisition. And the problem, it seems to me, arises primarily because cognitivists have tended to proceed as if the language-learner must *somehow* be formulating precise verbal hypotheses. It seems obvious, however, that far more fruitful parallels are to be found in the plethora of recent work on non-verbal elements in scientific discovery. Whether some version of Chomsky's argument for innate grammars can in the end be sustained, our appreciation of the issues can surely only be sharpened by attempting to view the language-learner as (*not* "unconsciously") developing non-verbal models, metaphors and analogies. The promise of this approach is that it gives us some idea of how light might dawn gradually over whole segments of his language as the subject progressively elaborates and fills in his intuitive models; and this gives reason for viewing scientific knowledge as parallel to linguistic understanding.

(3) In verbally representing the knowledge of someone who cannot give an accurate linguistic statement of it himself, it is often natural to employ

a notation whose conceptual power and precision go far beyond anything ascribable to the subject. Thus, in representing someone's knowledge of a melody it might be natural to employ conventional Western methods for specifying key- and time-signatures. Or again, in describing Newton's conception of light it might be natural to employ a battery of sophisticated mathematical treatments each of which is, on its own, far too precise, but which conjointly illuminate the question of what his view was. We are, indeed, familiar with this approach in philosophical exegesis. Thus, suppose we want to explain how Frege thought of reference. A natural procedure would be to give a selection of definitions, each one of which may exceed in clarity and precision Frege's own conception, but which nevertheless together shed light on his thought. The situation with regard to formal grammars and semantic theories is, I suggest, rather similar. They provide what is perhaps the most natural starting-point in attempting to articulate the structure of understanding; yet in clarity, precision and conceptual power they almost inevitably outstrip the grasp of language possessed by the general run of competent speakers.

This helps to explain why grammar and semantic theory often employ concepts not plausibly ascribable to competent speakers. Consider for example a central source of conceptual complexity in semantics: the use of quantifiers, which is responsible for talk of "satisfaction by infinite sequences". Most speakers do not possess the conception of satisfaction by an infinite sequence. Consider, however, an illustrative history of an individual's learning of quantification. We begin with quantification over finite domains ('All the toys in this box are broken') explained as equivalent to finite conjunction or disjunction ('This one is broken, and that one is, and . . .'). We step up to quantification over such domains as the natural numbers, again intuitively explained in terms of conjunction and disjunction; the analogy with the use already learnt is obvious, if not formally rigorous. On meeting the quantifier/variable notation, one is given the conception — and it has to be carefully explained — of objects realising the variables of open sentences. And eventually, one may meet with a Tarskian definition in terms of satisfaction by infinite sequences.

It is I think plausible to see these last two stages as involving a *crispening* of an intuitive model of what quantification means, which is *already* possessed by the subject. If this is correct, then surely there is no direct refutation of cognitivism on the basis of the appearance of talk of satisfaction by infinite sequences in semantic theory. It is simply a heuristically useful, but much too definite and precise, method of representing the knowledge which constitutes understanding of the natural-language quantifiers. The point is that we must not confuse truthlike insight into the nature of understanding with literal truth.

(4) There is, finally, one issue over knowledge-based approaches to language-mastery which may turn out to be the most pressing of all. It is raised by

a stimulating paper of J. M. Moravcsik [8], which aims to provide a general approach to many varieties of understanding, including for example comprehension of theories, substances, proofs and people. Moravcsik distinguishes “understanding what” and “understanding why” (which relate to, e.g., substances and events) from “understanding” *simpliciter* (which relates to, e.g., languages, proofs and people). He contends that the latter variety requires a “non-propositional conceptualisation” of the “essential structure” of the object of comprehension. Thus understanding of a language may require that one have some non-propositional grasp of its grammatical structure. And understanding a proof requires that one not only sees how each step follows from the preceding one, but that one can survey the structure of the proof.

The talk of “non-propositional conceptualisations” seems to come to this: it would be wrong to suppose that one could represent understanding, whether of a proof or of a language, as consisting *just* in knowledge *that* ϕ (for some appropriate ϕ). We may agree that the type of “grasp of structure” to which Moravcsik alludes may often be knowledge not readily verbalisable by those who possess it. It is often supposed — and the pressure to do so can seem overwhelming — that natural languages are, at least potentially, *universal*; in that either they, or extensions of them, can be used to *state* any possible knowledge that one might possess. The idea is that any *information* one might possess could be captured exactly by such a language. And Moravcsik, I suspect, would jib at this view. Certainly he would not be alone in doing so. Consider, for example, knowledge of someone’s appearance, gleaned say from observation of a photograph of them, so that one has an accurate mental image of how they look. Donald Davidson [2] writes: “How many facts or propositions are conveyed by a photograph? None, an infinity, or one great unstateable fact? Bad question. A picture is not worth a thousand words, or any other number. Words are the wrong currency to exchange for a picture” (p. 252). If this is correct, then knowledge of someone’s appearance may be something whose content cannot be verbally stated, and *a fortiori* cannot be represented by a sentence of the form ‘S knows that ϕ ’, where ‘ ϕ ’ is — of course — a sentence.

Without trying to do full justice to this line of thought, let me suggest that the cognitivist might simply take it in his stride. Even if the knowledge he aims to state cannot be captured *exactly* in language, he can still — verbally, of course, and within limits — make the situation clear. Why should knowledge not readily verbalisable by those who possess it resist verbal capture? Here are two possible reasons.

First, there is the *shifting* character of much of this type of knowledge. Thus I may know what *X* looks like: but perhaps sometimes I can think only of him as seen in profile, sometimes only as seen full-face, and perhaps never have a choice as to which. So a verbal statement of my knowledge might be excessively stable; perhaps inevitably so. Secondly, there is the

indefiniteness of much of this type of knowledge: as we have already seen, it is often natural to give a far too precise and conceptually powerful linguistic statement of what someone knows who cannot readily express the knowledge himself. The suggestion would be, now, that such excessive stability and precision is not only natural but inevitable: that language is too definite and sharp a medium ever to reflect accurately the delicately blurred nuances of non-linguistic thought.

To be sure, it is *we* who endow language with its semantic features; and so perhaps it is arguable that this train of thought is deeply incoherent; but I shall not examine the possibilities here. For even if something like this is in the end correct, the cognitivist can surely make the situation clear in developing his description of understanding. For suppose every representation that can be given of the knowledge constitutive of mastery is, when applied to any given speaker, inevitably too stable and precise. Then we can always *say* that it is; we can give stage-directions for the construal of the representation. And by giving a range of alternative accounts, all doubtless suffering from the same defects, but all equally good and all different, we may approximate ever more closely to his shifting, blurry knowledge in a stable, precise cognitive space.¹

Wolfson College, Oxford

BIBLIOGRAPHY

- [1] N. Chomsky, *Rules and Representations* (Oxford, 1980).
- [2] D. Davidson, "What Metaphors Mean", in *Reference, Truth and Reality*, ed. M. Platts (London, 1980), 238-54.
- [3] M. A. E. Dummett, "What is a Theory of Meaning?", in *Mind and Language*, ed. S. Guttenplan (Oxford, 1975), 97-138.
- [4] M. A. E. Dummett, "What is a Theory of Meaning? (II)", in *Truth and Meaning*, ed. G. Evans and J. McDowell (Oxford, 1976), 67-137.
- [5] M. A. E. Dummett, "What Does the Appeal to Use do for the Theory of Meaning?", in *Meaning and Use*, ed. A. Margalit (Dordrecht, 1979), 123-35.
- [6] G. Evans, "Semantic Theory and Tacit Knowledge", in *Wittgenstein: To Follow a Rule*, ed. S. Holtzman and C. H. Leich (London, 1981), 118-37.
- [7] G. Harman, "Language, Thought and Communication", in *Minnesota Studies in the Philosophy of Science*, vol. VII, ed. K. Gunderson (Minneapolis, 1975).
- [8] J. M. Moravcsik, "Understanding", *Dialectica*, 33 (1979), 201-16.
- [9] J. McDowell, "On the Sense and Reference of a Proper Name", *Mind*, 86 (1977), 159-68.
- [10] J. McDowell, "Physicalism and Primitive Denotation: Field on Tarski", in *Reference, Truth and Reality*, ed. M. Platts (London, 1980), 111-30.
- [11] F. Suppe, "The Search for Philosophic Understanding of Scientific Theories", in *The Structure of Scientific Theories*, ed. F. Suppe (Urbana, 1977).
- [12] C. Wright, "Rule-Following, Objectivity and the Theory of Meaning", in *Wittgenstein: To Follow a Rule*, ed. S. Holtzman and C. H. Leich (London, 1981), 99-117.

¹For criticism of earlier drafts, I am much indebted to Michael Dummett, John McDowell and Robert X. Ware.