

Mind, Language and Reality

Philosophical Papers, Volume 2

HILARY PUTNAM

Professor of Philosophy, Harvard University

NOTICE

This material may be
protected by copyright
law (Title 17 U.S. Code.)

NOTICE

This material may be
protected by copyright
law (Title 17 U.S. Code.)



CAMBRIDGE
UNIVERSITY PRESS

PUTNAM,

"IS SEMANTICS POSSIBLE?"

philosophers so grouped simply do not agree on sufficiently many doctrines or ideas for *one* attack to demolish all of them (if they can be demolished at all). Without committing myself to the rightness or wrongness of either Wittgenstein or Austin, I hope that I have succeeded in suggesting that these philosophers have far more intellectual and personal power than one could ever gather from such a description as the one Ayer gives.

Is semantics possible? *†

In the last decade enormous progress seems to have been made in the syntactic theory of natural languages, largely as a result of the work of linguists influenced by Noam Chomsky and Zellig Harris. Comparable progress seems *not* to have been made in the semantic theory of natural languages, and perhaps it is time to ask why this should be the case. Why is the theory of meaning so *hard*?

The meaning of common nouns

To get some idea of the difficulties, let us look at some of the problems that come up in connection with general names. General names are of many kinds. Some, like *bachelor*, admit of an explicit definition straight off ('man who has never been married'); but the overwhelming majority do not. Some are derived by transformations from verbal forms, e.g. *hunter* = *one who hunts*. An important class, philosophically as well as linguistically, is the class of general names associated with *natural kinds* – that is, with classes of things that we regard as of explanatory importance; classes whose normal distinguishing characteristics are 'held together' or even explained by deep-lying mechanisms. *Gold*, *lemon*, *tiger*, *acid*, are examples of such nouns. I want to begin this paper by suggesting that (1) *traditional* theories of meaning radically falsify the properties of such words; (2) logicians like Carnap do little more than formalize these traditional theories, inadequacies and all; (3) such semantic theories as that produced by Jerrold Katz and his co-workers likewise share all the defects of the traditional theory. In Austin's happy phrase, what we have been given by philosophers, logicians, and 'semantic theorists' alike, is a 'myth-eaten description'.

* Reprinted from H. Kiefer and M. Munitz (eds.) *Languages, Belief and Metaphysics*, Volume 1 of *Contemporary Philosophic Thought: The International Philosophy Year Conferences at Brockport* by permission of the State University of New York Press. Copyright © 1970 by State University of New York.

† While responsibility for the views expressed here is, of course, solely mine, they doubtless reflect the influence of two men who have profoundly affected my attitude towards the problems of language: Paul Ziff and Richard Boyd. I owe them both a debt of gratitude for their insight, their infectious enthusiasm, and for many happy hours of philosophical conversation.

In the traditional view, the meaning of, say 'lemon', is given by specifying a conjunction of *properties*. For each of these properties, the statement 'lemons have the property *P*' is an analytic truth; and if P_1, P_2, \dots, P_n are all the properties in the conjunction, then 'anything with all of the properties P_1, \dots, P_n is a lemon' is likewise an analytic truth.

In one sense, this is trivially correct. If we are allowed to invent unanalyzable properties *ad hoc*, then we can find a single property – not even a conjunction – the possession of which is a necessary and sufficient condition for being a lemon, or being gold, or whatever. Namely, we just postulate *the property of being a lemon*, or *the property of being gold*, or whatever may be needed. If we require that the properties P_1, P_2, \dots, P_n not be of this *ad hoc* character, however, then the situation is very different. Indeed, with any natural understanding of the term 'property', it is just *false* that to say that something belongs to a natural kind is just to ascribe to it a conjunction of properties.

To see why it is false, let us look at the term 'lemon'. The supposed 'defining characteristics' of lemons are: yellow color, tart taste, a certain kind of peel, etc. Why is the term 'lemon' *not* definable by simply conjoining these 'defining characteristics'?

The most obvious difficulty is that a natural kind may have *abnormal members*. A green lemon is still a lemon – even if, owing to some abnormality, it *never* turns yellow. A three-legged tiger is still a tiger. Gold in the gaseous state is still gold. It is only normal lemons that are yellow, tart, etc.; only normal tigers that are four-legged; only gold under normal conditions that is hard, white or yellow, etc.

To meet this difficulty, let us try the following definition: *X* is a *lemon* = *df*; *X* belongs to a natural kind whose normal members have yellow peel, tart taste, etc.

There is, of course, a problem with the 'etc.' There is also a problem with 'tart taste' – shouldn't it be *lemon* taste? But let us waive these difficulties, at least for the time being. Let us instead focus on the two notions that have come up with this attempted definition: the notions *natural kind* and *normal member*.

A natural kind *term* (so shift attention, for the moment, from natural kinds to their preferred designations) is a term that plays a special kind of role. If I describe something as a *lemon*, or as an *acid*, I indicate that it is likely to have certain characteristics (yellow peel, or sour taste in dilute water solution, as the case may be); but I also indicate that the presence of those characteristics, if they are present, is likely to be accounted for by some 'essential nature' which the thing shares with other members of the natural kind. What the essential nature is is not a

matter of language analysis but of scientific theory construction; today we would say it was chromosome structure, in the case of lemons, and being a proton-donor, in the case of acids. Thus it is tempting to say that a natural kind term is simply a term that plays a certain kind of role in scientific or pre-scientific theory: the role, roughly, of pointing to common 'essential features' or 'mechanisms' beyond and below the obvious 'distinguishing characteristics'. But this is vague, and likely to remain so. Meta-science is today in its infancy: and terms like 'natural kind', and 'normal member', are in the same boat as the more familiar meta-scientific terms 'theory' and 'explanation', as far as resisting a speedy and definitive analysis is concerned.

Even if we *could* define 'natural kind' – say, 'a natural kind is a class which is the extension of a term *P* which plays such-and-such a methodological role in some well-confirmed theory' – the definition would obviously embody a theory of the world, at least in part. It is not *analytic* that natural kinds are classes which play certain kinds of roles in theories; what *really* distinguishes the classes we count as natural kinds is itself a matter of (high level and very abstract) scientific investigation and not just meaning analysis.

That the proposed definition of 'lemon' uses terms which themselves resist definition is not a fatal objection however. Let us pause to note, therefore, that if it is correct (and we shall soon show that even it is radically oversimplified), then the traditional idea of the force of general terms is badly mistaken. To say that something is a lemon is, on the above definition, to say that it belongs to a natural kind whose normal members have certain properties; but not to say that it necessarily has those properties itself. There are no *analytic* truths of the form *every lemon has P*. What has happened is this: the traditional theory has taken an account which is correct for the 'one-criterion' concepts (i.e. for such concepts as 'bachelor' and 'vixen'), and made it a general account of the meaning of general names. A theory which correctly describes the behavior of perhaps three hundred words has been asserted to correctly describe the behavior of the tens of thousands of general names.

It is also important to note the following: if the above definition is correct, then knowledge of the properties that a thing has (in any natural and non 'ad hoc' sense of property) is not enough to determine, in any mechanical or algorithmic way, whether or not it is a lemon (or an acid, or whatever). For even if I have a description in, say, the language of particle physics, of what are in fact the chromosomal properties of a fruit, I may not be able to tell that it is a lemon because I have not developed the theory according to which (1) those physical-chemical

characteristics are the chromosomal structure-features (I may not even have the notion 'chromosome', and (2) I may not have discovered that chromosomal structure is the *essential* property of lemons. Meaning does not determine extension, in the sense that given the meaning and a list of all the 'properties' of a thing (in any particular sense of 'property') one can simply *read off* whether the thing is a lemon (or acid, or whatever). Even given the meaning, whether something is a lemon or not, is, or at least sometimes is, or at least may sometimes be, a matter of what is the best conceptual scheme, the best theory, the best scheme of 'natural kinds'. (This is, of course, one reason for the failure of phenomenalistic translation schemes.)

These consequences of the proposed definition are, I believe, correct, even though the proposed definition is itself still badly oversimplified. Is it a necessary truth that the 'normal' lemons, as we think of them (the tart yellow ones) are really normal members of their species? Is it logically impossible that we should have mistaken what are really very atypical lemons (perhaps diseased ones) for normal lemons? On the above definition, if there is no natural kind whose normal members are yellow, tart, etc., then even these tart, yellow, thick-peeled fruits that I make lemonade from are *not literally lemons*. But this is absurd. It is clear that they are lemons, although it is not analytic that they are *normal* lemons. Moreover, if the color of lemons changed – say, as the result of some gases getting into the earth's atmosphere and reacting with pigment in the peel of lemons – we would not say that lemons had ceased to exist, although a natural kind whose normal members were *yellow* and had the other characteristics of lemons *would* have ceased to exist. Thus the above definition is correct to the extent that what it says *isn't* analytic indeed *isn't*; but it is incorrect in that what would be analytic if it were correct *isn't*. We have loosened up the logic of the natural kind terms, in comparison with the 'conjunction of properties' model; but we have still not loosened it up enough.

Two cases have just been considered: (1) the normal members of the natural kind in question may not really be the ones we *think* are normal; (2) the characteristics of the natural kind may change with time, possibly due to a change in the conditions, without the 'essence' changing so much that we want to stop using the same word. In the first case (normal lemons are blue, but we haven't seen any normal lemons), our theory of the natural kind is false; but at least there is a natural kind about which we have a false theory, and that is why we can still apply the term. In the second case, our theory was at least once true; but it has ceased to be true, although the natural kind has not ceased to exist, which is why we can still apply the term.

Let us attempt to cover both these kinds of cases by modifying our definition as follows:

X is a *lemon* = *df* X belongs to a natural kind whose . . . (as before)
OR X belongs to a natural kind whose normal members used to . . . (as before) OR X belongs to a natural kind whose normal members were formerly believed to, or are now incorrectly believed to . . . (as before).

Nontechnically, the trouble with this 'definition' is that it is slightly crazy. Even if we waive the requirement of sanity (and, indeed, it is all too customary in philosophy to waive any such requirement), it still doesn't work. Suppose, for example, that some tens of thousands of years ago lemons were unknown, but a few atypical oranges were known. Suppose these atypical oranges had exactly the properties of peel, color, etc., that lemons have: indeed, we may suppose that only a biologist could tell that they were really queer oranges and not normal lemons. Suppose that the people living at that time took them to be normal members of a species, and thus thought that oranges have exactly the properties that lemons in fact do have. Then all now existing oranges would be lemons, according to the above definition, since they belong to a species (a natural kind) of which it was once believed that the normal members have the characteristics of yellow peel, lemon taste, etc.

Rather than try to complicate the definition still further, in the fashion of system-building philosophers, let us simply observe what has gone wrong. It is true – and this is what the new definition tries to reflect – that one possible use of a natural kind term is the following: to refer to a thing which belongs to a natural kind which does *not* fit the 'theory' associated with the natural kind term, but which was believed to fit that theory (and, in fact, to be *the* natural kind which fit the theory) when the theory had not yet been falsified. Even if cats turn out to be robots remotely controlled from Mars we will still call them 'cats'; even if it turns out that the stripes on tigers are painted on to deceive us, we will still call them 'tigers'; even if normal lemons are blue (we have been buying and raising very atypical lemons, but don't know it), they are still lemons (and so are the yellow ones.) Not only will we still *call* them 'cats', they are cats; not only will we still call them 'tigers', they are tigers; not only will we still call them 'lemons', they are lemons. But the fact that a term has several possible uses does not make it a disjunctive term; the mistake is in trying to represent the complex behavior of a natural kind word in something as simple as an analytic definition.

To say that an analytic definition is too simple a means of representation is not to say that no representation is possible. Indeed, a very simple representation is possible, namely:

lemon: natural kind word associated characteristics:
 yellow peel, tart taste, etc.

To fill this out, a lot more should be said about the linguistic behavior of natural kind words: but no more need be said about *lemon*.

Katz's theory of meaning

Carnap's view of meaning in natural language is this: we divide up logical space into 'logically possible worlds'. (That this may be highly language-relative, and that it may presuppose the very analytic-synthetic distinction he hopes to find by his quasi-operational procedure are objections he does not discuss.) The informant is asked whether or not he would say that something is the case in each logically possible world: the assumption being that (1) each logically possible world can be described clearly enough for the informant to tell; and (2) that the informant can say that the sentence in question is *true/false/not clearly either* just on the basis of the description of the logically possible world and the meaning (or 'intension') he assigns to the sentence in question. The latter assumption is false, as we have just seen, for just the reason that the traditional theory of meaning is false: even if I know the 'logically possible world' you have in mind, deciding whether or not something is, for example, a lemon, may require deciding what the best *theory* is; and this is not something to be determined by asking an informant yes/no questions in a rented office. This is not to say that 'lemon' has no meaning, of course: it is to say that meaning is not *that* simply connected with extension, even with 'extension in logically possible worlds'.

Carnap is not my main stalking-horse, however. The theory I want to focus on is the 'semantic theory' recently propounded by Jerrold Katz and his co-workers. In main outlines this theory is as follows:

(1) Each word has its meaning characterized by a string of 'semantic markers'.

(2) These markers stand for 'concepts' ('concepts' are themselves brain processes in Katz's philosophy of language; but I shall ignore this *jeu d'esprit* here). Examples of such concepts are: *unmarried*, *animate*, *seal*.

(3) Each such concept (concept for which a semantic marker is introduced) is a 'linguistic universal', and stands for an *innate* notion – one in some sense-or-other 'built into' the human brain.

(4) There are recursive rules – and this is the 'scientific' core of Katz's 'semantic theory' – whereby the 'readings' of whole sentences (these being likewise strings of markers) are derived from the meanings of the individual words and the deep structure (in the sense of transformational grammar) of the sentence.

(5) The scheme as a whole is said to be justified in what is said to be the manner of a scientific theory – by its ability to explain such things as our intuitions that certain sentences have more than one meaning, or the certain sentences are queer.

(6) Analyticity relations are also supposed to be able to be read off from the theory: for example, from the fact that the markers associated with 'unmarried' occur in connection with 'bachelor', one can see that 'all bachelors are unmarried' is analytic; and from the fact that the markers associated with 'animal' occur in connection with 'cat', one can see (allegedly) that 'all cats are animals' is analytic.

There are internal inconsistencies in this scheme which are apparent at once. For example, 'seal' is given as an example of a 'linguistic universal' (at least, 'seal' occurs as part of the 'distinguisher' in one reading for 'bachelor' – the variant reading: *young male fur seal*, in one of Katz's examples); but in no theory of human evolution is contact with seals universal. Indeed, even contact with *clothing*, or with *furniture*, or with *agriculture* is by no means universal. Thus we must take it that Katz means that whenever such terms occur they could be further analyzed into concepts which really are so primitive that a case could be made for their universality. Needless to say, this program has never been carried out, and he himself constantly ignores it in giving examples. But the point of greatest interest to us is that this scheme is an unsophisticated translation into 'mathematical' language of precisely the traditional theory that it has been our concern to criticize! Indeed, as far as general names are concerned, the only change is that whereas in the traditional account each general name was associated with a list of properties, in Katz's account each general name is associated with a list of *concepts*. It follows that each counterexample to the traditional theory is at once a counterexample also to Katz's theory. For example, if Katz lists the concept 'yellow' under the noun 'lemon', then he will be committed to 'all lemons are yellow'; if he lists the concept 'striped' under the noun 'tiger', then he will be committed to the analyticity of 'all tigers are striped'; and so on. Indeed, although Katz denies that his 'semantic markers' are themselves *words*, it is clear that they can be

regarded as a kind of artificial language. Therefore, what Katz is saying is that:

(1) A mechanical scheme can be given for translating any natural language into this artificial 'marker language' (and this scheme is just what Katz's 'semantic theory' is).

(2) The string of markers associated with a word has exactly the meaning of the word.

If (1) and (2) were true, we would at once deduce that there exists a possible language – a 'marker language' – with the property that every word that human beings have invented or could invent has an analytic definition in that language. But this is something that we have every reason to disbelieve! In fact: (1) We have just seen that if our account of 'natural kind' words is correct, then none of these words has an analytic definition. In particular, a natural kind word will be analytically translatable into marker language only in the special case in which a marker happens to have been introduced with its exact meaning. (2) There are many words for which we haven't the foggiest notion what an analytic definition would even look like. What would an analytic definition of 'mammoth' look like? (Would Katz say that it is analytic that mammoths are extinct? Or that they have a certain kind of molar? These are the items mentioned in the dictionary!) To say that a word is the name of an extinct species of elephant is to exactly communicate the use of that word; but it certainly isn't an analytic definition (i.e. an analytically necessary and sufficient condition). (3) *Theoretical terms* in science have no analytic definitions, for reasons familiar to every reader of recent philosophy of science; yet these are surely items (and not atypical items) in the vocabulary of natural languages.

We have now seen, I believe, one reason for the recent lack of progress in semantic theory: you may dress up traditional mistakes in modern dress by talking of 'recursive rules' and 'linguistic universals', but they remain the traditional mistakes. The problem in semantic theory is to get away from the picture of the meaning of a word as something like a *list of concepts*; not to formalize that misguided picture.

Quine's pessimism

Quine has long expressed a profound pessimism about the very possibility of such a subject as 'semantic theory'. Certainly we cannot assume that *there is* a scientific subject to be constructed here just because ordinary people have occasion to use the word 'meaning' from time to time; that would be like concluding that there must be a scientific subject to be constructed which will deal with 'causation' just

because ordinary people have occasion to use the word 'cause' from time to time. In one sense, *all* of science is a theory of causation; but not in the sense that it uses the word *cause*. Similarly, any successful and developed theory of language-use will in one sense be a theory of meaning; but not necessarily in the sense that it will employ any such notion as the 'meaning' of a word or of an utterance. Elementary as this point is, it seems to be constantly overlooked in the social sciences, and people seem constantly to expect that psychology, for example, must talk of 'dislike', 'attraction', 'belief', etc., simply because ordinary men use these words in psychological description.

Quine's pessimism cannot, then, be simply dismissed; and as far as the utility of the traditional notion of 'meaning' is concerned, Quine may well turn out to be right. But we are still left with the task of trying to say what are the real problems in the area of language-use, and of trying to erect a conceptual framework within which we can begin to try to solve them.

Let us return to our example of the natural kind words. It is a fact, and one whose importance to this subject I want to bring out, that the use of words can be taught. If someone does not know the meaning of 'lemon', I can somehow convey it to him. I am going to suggest that in this simple phenomenon lies the problem, and hence the *raison d'être*, of 'semantic theory'.

How do I convey the meaning of the word 'lemon'? Very likely, I show the man a lemon. Very well, let us change the example. How do I convey the meaning of the word 'tiger'? *I tell him what a tiger is.*

It is easy to see that Quine's own theoretical scheme (in *Word and Object*) will not handle this case very well. Quine's basic notion is the notion of *stimulus meaning* (roughly this is the set of nerve-ending stimulations which will 'prompt assent' to *tiger*). But: (1) it is very unlikely that I convey exactly the stimulus-meaning that 'tiger' has in my idiolect; and (2) in any case I don't convey it directly, i.e. by describing it. In fact, I couldn't describe it. Quine also works with the idea of *accepted sentences*; thus he might try to handle this case somewhat as follows: 'the hearer in your example already shares a great deal of language with you; otherwise you couldn't tell him what a tiger is. When you "tell him what a tiger is", you simply tell him certain sentences that you accept. Once he knows what sentences you accept, naturally he is able to use the word, at least observation words.'

Let us, however, refine this last counter somewhat. If conveying the meaning of the word 'tiger' involved conveying the totality of accepted scientific theory about tigers, or even the totality of what I believe about tigers, then it would be an impossible task. It is true that when I tell

someone what a tiger is I 'simply tell him certain sentences' – though not necessarily sentences I *accept*, except as descriptions of linguistically stereotypical tigers. But the point is, *which* sentences?

In the special case of such words as 'tiger' and 'lemon', we proposed an answer earlier in this paper. The answer runs as follows: there is somehow associated with the word 'tiger' a *theory*; not the actual theory we believe about tigers, which is very complex, but an oversimplified theory which describes a, so to speak, tiger *stereotype*. It describes, in the language we used earlier, a *normal member* of the natural kind. It is not necessary that we believe this theory, though in the case of 'tiger' we do. But it is necessary that we be aware that *this* theory is associated with the word: if our stereotype of a tiger ever changes, then the word 'tiger' will have changed its meaning. If, to change the example, lemons all turn blue, the word 'lemon' will not immediately change its meaning. When I first say, with surprise, 'lemons have all turned blue', lemon will still mean what it means now – which is to say that 'lemon' will still be associated with the stereotype *yellow lemon*, even though I will use the word to deny that lemons (even normal lemons) are in fact yellow. I can refer to a natural kind by a term which is 'loaded' with a theory which is known not to be any longer true of that natural kind, just because it will be clear to everyone that what I intend is to refer to *that* kind, and not to assert the theory. But, of course, if lemons really did turn blue (and stayed that way) then in time 'lemon' would come to have a meaning with the following representation:

lemon: natural kind word associated characteristics:
blue peel, tart taste, etc.

Then 'lemon' would have changed its meaning.

To sum this up: there are a few facts about 'lemon' or 'tiger' (I shall refer to them as *core facts*) such that one can convey the use of 'lemon' or 'tiger' by simply conveying those facts. More precisely, one cannot convey the approximate use *unless* one gets the core facts across.

Let me emphasize that this has the status of an empirical hypothesis. The hypothesis is that there are, in connection with almost any word (not just 'natural kind' words), certain core facts such that (1) one cannot convey the normal use of the word (to the satisfaction of native speakers) without conveying those core facts, and (2) in the case of many words and many speakers, conveying those core facts is sufficient to convey at least an approximation to the normal use. In the case of a natural kind word, the core facts are that a normal member of the kind has certain characteristics, or that this idea is at least the stereotype associated with the word.

If this hypothesis is false, then I think that Quine's pessimism is probably justified. But if this hypothesis is right, then I think it is clear what the problem of the theory of meaning is, regardless of whether or not one chooses to call it 'theory of *meaning*': the question is to explore and explain this empirical phenomenon. Questions which naturally arise are: what different kinds of words are associated with what different kinds of core facts? and by what mechanism does it happen that just conveying a small set of core facts brings it about that the hearer is able to imitate the normal use of a word?

Wittgensteinians, whose fondness for the expression 'form of life' appears to be directly proportional to its degree of preposterousness in a given context, say that acquiring the customary use of such a word as 'tiger' is coming to share a form of life. What they miss, or at any rate fail to emphasize, is that while the acquired disposition may be sufficiently complex and sufficiently interlinked with other complex dispositions to warrant special mention (though hardly the overblown phrase 'form of life'), what *triggers* the disposition is often highly discrete – e.g. a simple lexical definition frequently succeeds in conveying a pretty good idea of how a word is used. To be sure, as Wittgenstein emphasizes, this is only possible because we have a shared human nature, and because we have shared an acculturation process – there has to be a great deal of stage-setting before one can read a lexical definition and guess how a word is used. But in the process of 'debunking' this fact – the fact that something as simple as a lexical definition *can* convey the use of a word – they forget to be impressed by it. To be sure there is a great deal of stage-setting, but it is rarely stage-setting specifically designed to enable one to learn the use of *this* word. The fact that one *can* acquire the use of an indefinite number of new words, and on the basis of simple 'statements of what they mean', is an amazing fact: it is *the* fact, I repeat, on which semantic theory rests.

Sometimes it is said that the key problem in semantics is: how do we come to understand a new sentence? I would suggest that this is a far simpler (though not unimportant) problem. How logical words, for example, can be used to build up complex sentences out of simpler ones is easy to describe, at least in principle (of course, natural language analogues of logical words are far less tidy than the logical words of the mathematical logician), and it is also easy to say how the truth-conditions, etc., of the complex sentences are related to the truth-conditions of the sentences from which they were derived. This much *is* a matter of finding a structure of recursive rules with a suitable relation to the transformational grammar of the language in question. I would suggest that the question, How do we come to understand a new *word*? has far

more to do with the whole phenomenon of giving definitions and writing dictionaries than the former question. And it is this phenomenon – the phenomenon of writing (and needing) dictionaries – that gives rise to the whole idea of ‘semantic theory’.

Kinds of core facts

Let us now look a little more closely at the kind of information that one conveys when one conveys the meaning of a word. I have said that in the case of a ‘natural kind’ word one conveys the associated *stereotype*: the associated idea of the characteristics of a normal member of the kind. But this is not, in general, enough; one must also convey the extension, one must indicate *which* kind the stereotype is supposed to ‘fit’.

From the point of view of any traditional meaning theory, be it Plato’s or Frege’s or Carnap’s or Katz’s, this is just nonsense. How can I ‘convey’ the extension of, say ‘tiger’? Am I supposed to give you all the tigers in the world (heaven forbid!). I can convey the extension of a term only by giving a description of that extension; and then that description must be a ‘part of the meaning’, or else my definition will not be a meaning-statement at all. To say: ‘I gave him certain conditions associated with the word, *and* I gave him the extension’ (as if that weren’t just giving *further* conditions) can only be nonsense.

The mistake of the traditional theorist lies in his attachment to the word ‘meaning’. If giving the meaning is *giving* the *meaning*, then it is giving a definite thing; but giving the meaning isn’t, as we shall see in a moment, giving some one definite thing. To drop the word ‘meaning’, which is here extremely misleading: there is no *one* set of facts which has to be conveyed to convey the normal use of a word; and taking account of this requires a complication in our notion of ‘core facts’.

That the same stereotype might be associated with different kinds seems odd if the kind word one has in mind is ‘tiger’; but change the example to, say, ‘aluminum’ and it will not seem odd at all. About all I know about aluminum is that it is a light metal, that it makes durable pots and pans, and that it doesn’t appear to rust (although it does occasionally discolor). For all I know, every one of these characteristics may also fit molybdenum.

Suppose now that a colony of English-speaking Earthlings is leaving in a spaceship for a distant planet. When they arrive on their distant planet, they discover that no one remembers the atomic weight (or any other defining characteristic) of aluminum, nor the atomic weight (or other characteristic) of molybdenum. There is some aluminum in the spacecraft, and some molybdenum. Let us suppose that they guess

which is which, and they guess wrong. Henceforth, they use ‘aluminum’ as the name for molybdenum, and ‘molybdenum’ as the name for aluminum. It is clear that ‘aluminum’ has a different meaning in this community than in ours: in fact, it means *molybdenum*. Yet how can this be? Didn’t they possess the normal ‘linguistic competence’? Didn’t they all ‘know the meaning of the word “aluminum”’?

Let us duck this question for a moment. If I want to make sure that the word ‘aluminum’ will continue to be used in what counts as a ‘normal’ way by the colonists in my example, it will suffice to give them some test for aluminum (or just to give them a carefully labelled sample, and let them discover a test, if they are clever enough). Once they know how to *tell* aluminum from other metals, they will go on using the word with the correct extension as well as the correct ‘intension’ (i.e. the correct stereotype). But notice: it does not matter *which* test we give the colonists. The test isn’t part of the meaning; but that there be some test or other (or something, e.g. a sample, from which one might be derived), is necessary to preservation of ‘the normal usage’. Meaning indeed determines extension; but only because extension (fixed by *some* test or other) is, in some cases, ‘part of the meaning’.

There are two further refinements here: if we give them a test, they mustn’t make it part of the stereotype – that would be a change of meaning. (Thus it’s better if they don’t all *know* the test; as long as only experts do, and the average speaker ‘asks an expert’ in case of doubt, the criteria mentioned in the test can’t infect the stereotype.) Asking an expert is enough of a test for the normal speaker; that’s why we don’t give a test in an ordinary context.

We can now modify our account of the ‘core facts’ in the case of a natural kind word as follows: (1) The core facts are the stereotype *and* the *extension*. (2) Nothing normally need be said about the extension, however, since the hearer knows that he can always consult an expert if any question comes up. (3) In special cases – such as the case of colonists – there may be danger that the word will get attached to the wrong natural kind, even though the right stereotype is associated with it. In such cases, one must give some way of getting the extension right, but no one *particular* way is necessary.

In the case of ‘lemon’ or ‘tiger’ a similar problem comes up. It is logically possible (although empirically unlikely, perhaps) that a species of fruit biologically unrelated to lemons might be indistinguishable from lemons in taste and appearance. In such a case, there would be two possibilities: (1) to call them *lemons*, and thus let ‘lemon’ be a word for any one of a number of natural kinds; or (2) to say that they are not

lemons (which is what, I suspect, biologists would decide to do). In the latter case, the problems are exactly the same as with *aluminum*: to be sure one has the 'normal usage' or 'customary meaning' or whatever, one has to be sure one has the right extension.

The problem: that giving the extension is part of giving the meaning arises also in the case of names of sensible qualities, e.g. colors. Here, however, it is normal to give the extension by giving a sample, so that the person learning the word learns to recognize the quality in the normal way. Frequently it has been regarded as a defect of *dictionaries* that they are 'cluttered up' with color samples, and with stray pieces of empirical information (e.g. the atomic weight of aluminum), not sharply distinguished from 'purely linguistic' information. The burden of the present discussion is that this is no defect at all, but essential to the function of conveying the core facts in each case.

Still other kinds of words may be mentioned in passing. In the case of 'one-criterion' words (words which possess an analytical necessary and sufficient condition) it is obvious why the core fact is just the analytical necessary and sufficient condition, e.g. 'man who has never been married', in the case of 'bachelor'. In the case of 'cluster' words (e.g. the name of a disease which is known not to have any one underlying cause), it is obvious why the core facts are just the typical symptoms or elements of the cluster; and so on. Given the *function* of a kind of word, it is not difficult to explain why certain facts function as core facts for conveying the use of words of that kind.

The possibility of semantics

Why, then, is semantics so hard? In terms of the foregoing, I want to suggest that semantics is a typical social science. The sloppiness, the lack of precise theories and laws, the lack of mathematical rigor, are all characteristic of the social sciences today. A general and precise theory which answers the questions (1) why do words have the different sorts of functions they do? and (2) exactly how does conveying core facts enable one to learn the use of a word? is not to be expected until one has a general and precise model of a language-user; and that is still a long way off. But the fact that Utopia is a long way off does not mean that daily life should come to a screeching halt. There is plenty for us to investigate, in our sloppy and impressionistic fashion, and there are plenty of real results to be obtained. The first step is to free ourselves from the oversimplifications foisted upon us by the tradition, and to see where the real problems lie. I hope this paper has been a contribution to that first step.

The refutation of conventionalism*

I shall discuss conventionalism in Quine's writings on the topic of radical translation, and in the writings of Reichenbach and Grünbaum on the nature of geometry.

Let me say at the outset that Quine and Reichenbach are the two philosophers who have had the greatest influence on my own philosophical work. If Quine's ideas have not had the full influence they deserve, it may be in part because of the intensely paradoxical nature of the doctrines put forward – or seemingly put forward – in *Word and Object*. The doctrines of *Word and Object* – in particular the impossibility of radical translation – look wrong to many philosophers. Since these doctrines are thought by Quine himself to follow from the doctrines put forward in 'Two dogmas of empiricism', they cast doubt on 'Two dogmas' itself. My contention here will be that the impossibility of radical translation does not follow from the critique of the analytic-synthetic distinction. I believe that Quine is right in his critique of the analytic-synthetic distinction, but wrong in his argument for the impossibility of radical translation. By showing that one does not have to accept Quine's arguments for the impossibility of radical translation, I hope, therefore, to clarify just what it is in Quine's work that is true and important.

Similarly, I think that Reichenbach understood the importance of non-Euclidean geometry for epistemology as it has not been understood by philosophers in general to the present day. I think that he understood the issues surrounding the verifiability theory of meaning in his book *Experience and Prediction* far more deeply than they have been generally understood, and I think his writings on the philosophical foundations of induction clarify the nature of the problems in an unparalleled way, even if Reichenbach's particular solutions cannot be accepted. In Reichenbach's case too, it is a peculiar kind of conventionalism that seems to flow from his work that has led philosophers to downgrade him or even to dismiss him as just a 'positivist'. Again, I

* First published in M. Munitz (ed.) *Semantics and Meaning* (New York, 1975). Reprinted by permission of New York University Press.