

beyond the most notorious examples, and that perhaps not the whole of this range would prove amenable to the envisaged mode of treatment. At this point I took up the idea of a minimal strengthening of a Russellian pattern of analysis by the addition of a purely syntactical scope device, which could at the same time be regarded as a conventional regimentation of a particular kind of nonconventional implicature. I have not, in this essay, given any consideration at all to what might well turn out to be the best treatment of definite descriptions, namely to the idea that they are, in the first instance at least, to be regarded as being, semantically, a special subclass of referential expressions.

THIS MATERIAL MAY BE  
PROTECTED BY  
COPYRIGHT LAW  
(TITLE 17, U.S. CODE)

18

## Meaning Revisited

I am going to treat informally some topics connected with meaning. Rather than trying to say something particularly new, I have the idea of putting one or two of the thoughts I have had at various times into some kind of focus, so that there might emerge some sense about not merely what kind of view about the nature of meaning I am, or was, inclined to endorse, but also why it should be antecedently plausible to accept this kind of view. When I say "antecedently plausible," I mean plausible for some reasons other than that the view in question offers some prospects of dealing with the intuitive data: the facts about how we use the word "mean," and so on. So I shall be digging just a little bit into the background of the study of meaning and its roots in such things as philosophical psychology; but I hope without any very formidable detail.

The main theme will be matters connected with the relation between speaker's meaning and meaning in a language, or word meaning, sentence meaning, expression meaning, and so on. In the course of this account, I shall make reference to something like the definitions or analyses that I have previously offered; I do not guarantee that any that I use will be quite the same, but this does not worry me because I am not here concerned with the details. It seems to me that with regard to the possibility of using the notion of intention in a nested kind of way to explicate the notion of meaning, there are quite a variety of plausible, or at least not too implausible, analyses, which differ to a greater or lesser extent in detail, and at the moment I am not really concerned with trying to adjudicate between the various versions.

The essay will be in three sections. In the first, I shall try to sketch three kinds of correspondence which one might be justified in looking for, or even demanding, when thinking about thought, the world, and language. I hope this will provide some sort of a framework within which to set views about meaning; it may in fact offer some sort of impetus toward this or that view. In the second section, which will again involve an attempt to fit things into a wider framework, I shall provide some discussion of what I once called the distinction between natural and nonnatural meaning. Here I am interested not so much in the existence of that distinction, which has now, I think, become pretty boringly common ground (or mutual knowledge), but rather in the relationship between the two notions, the connections rather than the dissimilarities between them. I shall announce the contents of the third section when I come to it; we'll keep it as a mystery package.

### I. Language, Thought, and Reality

The first of the three correspondences which one might expect to find when thinking in largish terms about the relationships between reality, thought, and language or communication devices, is a correspondence between thought and reality: what I shall call, for some kind of brevity, psychophysical correspondence. This has obvious connections with the general idea of truth in application to beliefs or analogous notions describing physical states. The point I want to get at is that it is not just that there *are* such correspondences, or that it is intuitively plausible to think there are, but rather that their presence is *needed*, or *desirable*, if one looks at the ways in which human beings and other sentient creatures get around and stay alive, as well as perhaps doing more ambitious things than that. This leads to a view, which I have held for some time, of a battery of psychological concepts which we use both about ourselves and about what one might think of as lower creatures, as having the function of providing an explanatory bridge between the appearance of a creature in a certain kind of physical situation and its engaging in certain sorts of behavior.

For instance, suppose we have a creature C that is in the presence of a certain object, let's say a piece of cheese, and we get a situation in which the creature eats the object. In certain circumstances, we might want to invoke the contents of a psychological theory in order

to explain the transition from the creature's being in the presence of the object to its eating it. The "bridgework" which would be done in these terms can be put in a rather schematic and unrealistic way, which again does not worry me since I am not trying to provide a proper explanation, but merely a rough idea of what the pattern of a proper explanation would be.

First, let us suppose that the creature believes (or thinks, if it is not advanced enough to have beliefs) that the object is a bit of cheese, and that it also believes that the object is nearby. Second, let us suppose that the creature believes, or thinks, that cheese is something to eat; and third, that the creature is hungry, that it wants to eat. All this looks relatively unexciting, just as it should. Then by virtue of what I would think of as a vulgar, vernacular, psychological law, the operation of which is the reason for the introduction of the concepts of believing and wanting, we get our first psychological law for the creature or type of creature in question. This is that for any particular object X and for any feature F and for any activity or type of behavior A, if the creature C believes that the object X both has the feature F and is nearby, or within reach, and that things of type F are suitable for activity A, then the creature wants to A with respect to the object X. In other words, the law harnesses the object to the type of activity.

Applying the law to the three initial premises given above, we reach a further stage: namely, that the creature C wants to eat the object X, that is, the piece of cheese. We might then invoke a second psychological law for this creature, this time a psychophysical law: that for any type of activity A, if a creature C, wants to A with respect to some particular object X, and if it is not prevented in one or another of a set of ways which might or might not be listable, then creature C does produce the activity A with respect to X. And with the application of the second law, we get to our final step, which is that the creature C eats the object X. We now have our explanation, a bridge between the initial situation, the creature's being in the presence of a piece of cheese, and the final behavior, the creature's eating it.

The laws I have mentioned are vulgar laws. The kind of theory in which I think of them appearing would not be a specialist or formalized psychological theory, if indeed there are such things; I am perhaps not very comfortable with the word "theory" being applied to it. It would be the rough kind of system with which we all work, and the laws in it are to be thought of as corrigible, modifiable and *ceteris paribus* in character.

Now the creature C may be frustrated if certain psychophysical correspondences do not obtain. For instance, if C believes *wrongly* that the object in front of it is a piece of cheese, or thinks *wrongly* of cheese as being from its point of view something to eat, then at the very least, C may get indigestion when it consumes the object. For this reason, psychophysical correspondences are required (things like beliefs have to be true, and so on) for the operation of the psychological mechanisms which I have sketched to be *beneficial* to the creature in question.

In a similar way, if the creature's desire to eat the object is for some reason not fulfilled, the creature stays hungry. Here again we need some kind of correspondences, parallel to those between beliefs and reality, between desires, wants, states of will and so on, and reality. Desires, etc. need to correspond to reality in order to be fulfilled: that is, in order for the psychological mechanism to operate in a beneficial way.

Finally, if C is a rational or reflective creature (which I have not so far been assuming), it may recognize the kind of facts about itself that I have just sketched, may recognize that correspondences between psychological states and the world are in general required for the psychological mechanism to be useful, and may also recognize that subject to this proviso of correspondence, the psychological mechanism is conducive to survival or to the attainment of other of its objectives. If it recognizes all this, then it will presumably itself think of such correspondences as being desirable things to have around from its own point of view: that is, the correspondences will not only *be* desirable but will be regarded by C as desirable.

In the first place, then, we have psychophysical correspondences, and they seem to be the kind of things that one would want to have. But there are further correspondences too. There is, as it were, a triangle consisting of reality, thought, and language or communication devices, and we now have, I hope, one hookup, between thought and reality. However, I think there is also a hookup of a different kind, which would also be desirable, between thought and communication devices. This would again involve correspondences of a relatively simple and obvious kind, along the following lines.

First, the operation of such creatures as I have been talking about is at least in certain circumstances going to be helped and furthered if there is what one might think of as shared experience. In particular, if psychological states which initially attach to one creature can be

transmitted or transferred or reproduced in another creature (a process which might be called  $\psi$ -transmission), that would be advantageous. Obviously, the production of communication devices is a resource which will help to effect such transfers.

If one accepts this idea, then one could simply accept that for the process to be intelligible, understandable, there will have to be correspondences between particular communication devices or utterances on the one hand, and psychological states on the other. These correspondences may be achieved either directly, or (more likely) indirectly, via the *types* to which the particular utterances belong: the sentences which the particular utterances are utterances of, the gesture-types which the particular gestures are productions of, and so on. Whether direct or indirect, the correspondence would be between utterances or utterance-types on the one hand, and *types* of psychological states on the other, where these would include, for example, the belief-types to which the beliefs of particular people belong: not *Jones's* belief that such-and-such, but *a* belief that such-and-such.

If there exist these correspondences between utterances or utterance-types on the one hand, and psychological types on the other, we can say that it is in general, and subject to certain conditions, *desirable* for there to occur, in the joint or social lives of creatures of the kind in question, sequences of the following sort: a certain psychological state  $\psi^1$  in certain circumstances is followed by a certain utterance U, made in certain circumstances, which in turn, if the circumstances are right, is followed by a particular instance of a further psychological type  $\psi^2$ , a state not now in the communicating creature but in the creature who is communicated to. And it might be a matter of desirability for  $\psi^1$  and  $\psi^2$  to be states of one and the same, rather than of different sorts, so that when these sequences  $\psi^1$ , U,  $\psi^2$  occur, they involve utterances and psychological states between which these psycholinguistic correspondences obtain.

Of course, transfers can occur without these correspondences obtaining. A creature may choose the wrong utterance to express its psychological state, and then there will probably be a misfire in the utterance, and the psychological state induced in the second creature will not be the same as the one present in the first. Alternatively, the first creature may operate all right, but the second creature may, as it were, misunderstand the device that was produced, and so pick up the wrong belief or desire, one that does not correspond to the utterance produced by the first creature.

The general condition, or at least the most salient general condition, for the desirability or beneficial character of transfers of this sort is connected with the obtaining of the first kind of correspondence. That is, if all these transfers were to involve the transmission of mistaken beliefs, it is not clear that one would regard this communication mechanism as beneficial, even if the appropriate psycholinguistic correspondences held. I do think it is in some sense inconceivable that all the transfers should involve mistaken beliefs, but at least such a state of affairs can be contemplated. So a general condition would be that soul-to-soul transfers, so to speak, are beneficial provided that the states transmitted are ones which correspond with the world.

It looks now as if we have got to a point at which we have in outline, presented in a sort of general, semitheological way, a rough prototype of a notion of truth in application to beliefs and such-like things. That would give the first kind of correspondence. In the second kind of correspondence, we have what is at least a promising candidate for being a rough prototype of the notion of meaning, for it looks as though it is not implausible to suggest that to explain, with respect to some particular utterance or utterance-type what type of psychological state it corresponds to, in such a way that transfers of this kind are characteristically a feature of creatures' lives, would be a first approximation to explaining the meaning of the utterances or utterance-types in question. We have thus hooked up all three corners: corner number one, reality, has been hooked up with corner number two, thought, and we have hooked up thought, corner number two, with language or communication devices, corner number three. This of course yields a derivative link between corner number three, utterances or sentences, and corner number one, reality, via the beliefs or other psychological states with which they are themselves connected. But it is perhaps also worthwhile to ask whether in addition to such an indirect connection between language and reality, one which proceeds through the intermediation of psychological states, there is an arguable possibility of a direct link: a direct line between language and reality as well as a line through thought. And I think it is at least arguable that there is, though of course one of a kind which we will have to harmonize with the links that have already been introduced.

If I ask the question "What are the conditions for a specified belief (for instance a belief that snow is white) corresponding with the world?" I can give answers for individual cases without much diffi-

culty. For instance, I can say that a belief that snow is white will correspond to the world just in case snow is white, and I can say that a belief that cheese is blue will correspond to the world just in case cheese is blue; and you can see perfectly well what the rather dreary routine is that I use in these particular cases. However, it has been noticed by philosophers that there are difficulties about explicitly generalizing the individual bits of communication of which I have just spouted one or two samples. That is because to generalize from them would presumably be to omit reference to the particular objects and the particular beliefs, and to state in general what the conditions are for beliefs corresponding to objects. That is to say, a general condition of correspondence between beliefs and the world would have to begin something like: "For any item which one believes, that item corresponds with the world if . . ."; but then how does one go on? In the particular cases, I had a sentence which I had cited or referred to in the antecedent on the left-hand side, which I then produced on the right-hand side; but since I have eliminated all reference to particular beliefs or sentences, I no longer have a sentence available with which to complete the general condition.

It looks as if I should want to say something like: "For any item P, if one believes that P, then one's belief that P corresponds with the world just in case P." Unfortunately that involves difficulties, because by the ordinary account of quantification, I am talking about objects or items, so I might just as well use the letter x, characteristic of objects, and say: "For any item x, if x is believed, then x corresponds with the world just in case x." But that seems rather like producing the generalization that if any object x is a pig, then x: and that is not an intelligible form of statement, because "x" is not a variable for which one can substitute sentences. In fact, something seems to have got left out somewhere, and we have not got an intelligible specification of truth-conditions. Moreover, this is difficult to remedy, because without getting into tortures over shifts between thinking of propositions or propositional expressions as sometimes like names (with the form of *that*-clauses) and sometimes like sentences (the result of detaching the word "that" from these *that*-clauses), it is difficult to know what to do.

It looks as though, to avoid this difficulty, if I want to produce a generalization of the idea of correspondences between psychological states and the world which I have already in some sense provided for, I might well have to use some form like the following: "If a certain

sentence S is an expression of a certain particular belief, then the belief which the sentence S expresses corresponds with the world just in case S is true." That is to say, to remedy this difficulty in generalization, I now bring into play a notion of truth in application to sentences, and I do this in order to have a way of stating generally the conditions for the correspondence of beliefs and the world. Thus, in trying to safeguard the characterization of what it is for beliefs to correspond with the world, I have introduced another correspondence, a correspondence between utterances or sentences and the world, signalized by the appearance of the word "true."

It looks, then, as if in order to achieve a characterization of the first kind of correspondence, between beliefs and the world, one has to make use of a parallel kind of correspondence between utterances or sentences and the world. Hence these latter correspondences may be not only possible but needed if one is to be able to state, in a general way, that correspondences of the psychophysical kind actually obtain.

However, though they may be required for expository purposes, as I have sketched, it might still be the case that in order to show that correspondences between sentences and the world were desirable, not just for purposes of theoretical exposition but from the point of view of creatures who operate with such utterances or utterance-types, one still has to bring in the psychological states in specifying the conditions of suitability, desirability, or whatever. That is, it might be that one can certainly formulate or characterize some notion of direct correspondence between utterances and the world, and this might have a certain limited teleological justification because it is needed to provide a general way of expressing the conditions for other types of correspondence, but if one wants to provide a more general teleological justification, one would need to make reference to beliefs and other psychological states. In other words, for a more general justification of the idea of truth in application to sentences, one might have to bring in all three corners, including the missing one.

## II. Natural and Nonnatural Meaning

I have now, as it were, smuggled in some sort of preliminary version of the kind of view about meaning which I have gone on record as holding; but that is qualified by the fact that I have also smuggled in versions of the kind of views that other people have gone in for too. Let me now try to advance the case a little further, in a way which

might support my sort of view as opposed to certain other views.

I do not think it is too controversial to advance the idea that there is a reasonably clear intuitive distinction between cases where the word "mean" has what one might think of as a natural sense, a sense in which what something means is closely related to the idea of what it is a natural sign for (as in "Black clouds mean rain"), and those where it has what I call a nonnatural sense, as in such contexts as "His remark meant so-and-so," "His gesture meant that he was fed up," and so on (cf. Essay 14).

I have offered one or two recognition tests which might enable one to tell which of these, natural or nonnatural meaning, one was actually dealing with in a given case. The tests were, roughly speaking, that the nonnatural cases of meaning, cases which are related to communication, are what we might call nonfactive, whereas the natural cases are factive. That is, anyone who says "Those black clouds mean rain," or "Those black clouds meant that it would rain," would presumably be committing himself to its being the case that it will rain, or that it did rain. However, if I say "His gesture meant that he was fed up," under an interpretation of a nonnatural kind, one specially connected with what we think of as communication, then to say that does not commit you to his actually being fed up. I also noted that the specification of the nonnatural meaning of items can be comfortably done via the use of phrases in quotation marks, whereas it would seem rather odd to say that those black clouds meant "It will rain": it does not look as if one can replace the *that*-clause here by a sentence in quotation marks.

Assuming for the moment that these tests are roughly adequate, what I want to do now is not to emphasize the differences between these cases, because that has already been done, but rather to look at what they have in common. Is this double use of the word "mean" just like the double use of the word "vice" to refer sometimes to something approximating to a sin and sometimes to a certain sort of instrument used by carpenters? One is pretty much inclined in the latter case to say that there are two words which are pronounced and written the same.

On general grounds of economy, I am inclined to think that if one can avoid saying that the word so-and-so has this sense, that sense, and the other sense, or this meaning and another meaning, if one can allow them to be variants under a single principle, that is the desirable thing to do: don't multiply senses beyond necessity. And it occurs to

me that the root idea in the notion of meaning, which in one form or adaptation or another would apply to both of these cases, is that if  $x$  means that  $y$ , then this is equivalent to, or at least contains as a part of what it means, the claim that  $y$  is a consequence of  $x$ . That is, what the cases of natural and nonnatural meaning have in common is that, on some interpretation of the notion of consequence,  $y$ 's being the case is a consequence of  $x$ .

Of course one will expect there to be differences in the kind of consequence involved, or the way in which the consequence is reached. So what I want to do now is look to see if one would represent the cases of nonnatural meaning as being descendants from, in a sense of "descendant" which would suggest that they were derivative from and analogous to, cases of natural meaning. I shall also look a little at what kind of principles or assumptions one would have to make if one were trying to set up this position that natural meaning is in some specifiable way the ancestor of nonnatural meaning.

In the case of natural meaning, among the things which have natural meaning, besides black clouds, spots on the face, and symptoms of this or that disease, are certainly forms of behavior: things like groans, screeches, and so on, which mean, or normally mean, that someone or something is in pain or some other state. Thus special cases of natural meaning are cases in which bits of things like bodily behavior mean the presence of various elements or states of the creature that produces them. In the natural case, the production of these pieces of behavior, or at least the presence in those pieces of behavior of the particular features which, as it were, do the meaning for one, is nonvoluntary. Thus we have as a sort of canonical pattern that some creature  $X$  nonvoluntarily produces a certain piece of behavior  $\alpha$ , the production of which means, or has the consequence, or evidences, that  $X$  is in pain. That is the initial natural case. Let us now see if we could in one or more stages modify it so as to end up with something which is very much like nonnatural meaning.

Stage one in the operation involves the supposition that the creature actually voluntarily produces a certain sort of behavior which is such that its nonvoluntary production would be evidence that the creature is, let us say, in pain. The kinds of cases of this which come most obviously to mind will be cases of faking or deception. A creature normally voluntarily produces behavior not only when, but *because*, its nonvoluntary production would be evidence that the creature is in a certain state, with the effect that the rest of the world,

other creatures around, treat the production, which is in fact voluntary, as if it were a nonvoluntary production. That is, they come to just the same conclusion about the creature's being in the state in question, the signaled state. The purpose of the creature's producing the behavior voluntarily would be so that the rest of the world should think that it is in the state which the nonvoluntary production would signify.

In stage two not only does creature  $X$  produce this behavior voluntarily instead of nonvoluntarily, as in the primitive state, but we also assume that it is *recognized* by another creature  $Y$ , involved with  $X$  in some transaction, as being the voluntary production of a certain form of behavior the nonvoluntary production of which evidences, say, pain. That is, creature  $X$  is now supposed not only to simulate pain-behavior, but also to be recognized as simulating pain-behavior. The import of the recognition by  $Y$  that the production is voluntary undermines, of course, any tendency on the part of  $Y$  to come to the conclusion that creature  $X$  is in pain. So, one might ask, what would be required to restore the situation: what could be added which would be an antidote, so to speak, to the dissolution on the part of  $Y$  of the idea that  $X$  is in pain?

A first step in this direction would be to go to what we might think of as stage three. Here, we suppose that creature  $Y$  not only recognizes that the behavior is voluntary on the part of  $X$ , but also recognizes that  $X$  *intends*  $Y$  to recognize his behavior as voluntary. That is, we have now undermined the idea that this is a straightforward piece of deception. Deceiving consists in trying to get a creature to accept certain things as signs of something or other without knowing that this is a faked case. Here, however, we would have a sort of perverse faked case, in which something is faked but at the same time a clear indication is put in that the faking has been done.

Creature  $Y$  can be thought of as initially baffled by this conflicting performance. There is this creature, as it were, simulating pain, but announcing, in a certain sense, that this is what it is doing: what on earth can it be up to? It seems to me that if  $Y$  does raise the question of why  $X$  should be doing this, it might first come up with the idea that  $X$  is engaging in some form of play or make-believe, a game to which, since  $X$ 's behavior is seemingly directed toward  $Y$ ,  $Y$  is expected or intended to make some appropriate contribution. Cases susceptible of such an interpretation I regard as belonging to stage four.

But, we may suppose, there might be cases which could not be handled in this way. If Y is to be expected to be a fellow-participant with X in some form of play, it ought to be possible for Y to recognize what kind of contribution Y is supposed to make; and we can envisage the possibility that Y has no clue on which to base such recognition, or again that though some form of contribution seems to be suggested, when Y obliges by coming up with it, X, instead of producing further play-behavior, gets cross and perhaps repeats its original, and now problematic, performance.

We now reach stage five, at which Y supposes not that X is engaged in play, but that what X is doing is trying to get Y to believe or accept that X is in pain: that is, trying to get Y to believe in or accept the presence of that state in X which the produced behavior, when produced nonvoluntarily, is in fact a natural sign of, naturally means. More specifically, one might say that at stage five creature Y recognizes that creature X in the first place intends that Y recognize the production of the sign of pain (of what is usually the sign of pain) to be voluntary, and further intends that Y should regard this first intention as being a sufficient reason for Y to believe that X is in pain; and that X has these intentions because he has the additional further intention that Y should not merely have sufficient reason for believing that X is in pain, but should actually believe it.

Whether or not in these circumstances Y will not merely recognize that X intends, in a certain rather queer way, to get Y to believe that X is in pain, whether Y not only recognizes this but actually goes on to believe that X is in pain, would presumably depend on a further set of conditions which can be summed up under the general heading that Y should regard X as trustworthy in one or another of perhaps a variety of ways. For example, suppose Y thinks that, either in general or at least in this type of case, X would not want to get Y to believe that X is in pain unless X really were in pain. Suppose also (this would perhaps not apply to a case of pain but might apply to the communication of other states) that Y also believes that X is trustworthy, not just in the sense of not being malignant, but also in the sense of being, as it were, in general responsible, for example, being the sort of creature who takes adequate trouble to make sure that what he is trying to get the other creature to believe is in fact the case, and who is not careless, negligent, or rash. Then, given the general fulfillment of the idea that Y regards X either in general or in this particular case as being trustworthy in this kind of competent, careful

way, one would regard it as rational not only for Y to recognize these intentions on the part of X, that Y should have certain beliefs about X's being in pain, but also for Y actually to pass to adopting these beliefs.

So far I have been talking about the communication of the idea that something is the case, for example, that X is in pain or some other state, by means of a nondeceptive simulation on the part of the communicating creature of the standard signs or indices of such a state. But the mechanism that has been used, involving the interchange of beliefs or intentions of different orders, really does not require that what is taken as the communication vehicle should be initially a natural expression or sign of the state of affairs being communicated. If we now relax this requirement, we get to stage six, the road to which may be eased by the following reflection, for which I am indebted to Judith Baker.

In relation to the particular example which I have been using, to reach the position ascribed to it in stage five, Y would have to solve, bypass, or ignore a possible problem presented by X's behavior: why should X produce what is not a genuine but a faked expression of pain if what X is trying to get Y to believe is that X is *in pain*? Why not just let out a natural bellow? Possible answers are not too hard to come by: for example, that it would be unmanly, or otherwise uncreatively, for X to produce *naturally* a natural expression of pain, or that X's nonnatural production of an expression of pain is not to be supposed to indicate *every* feature which would be indicated by a natural production (the nonnatural emission, for example, of a loud bellow might properly be taken to indicate pain, but not that degree of pain which would correspond with the decibels of the particular emission). This problem would not, however, arise if X's performance, instead of being something which, in the natural case, would be an *expression* of that state of X which (in the nonnatural case) it is intended to get Y to believe in, were rather something more loosely connected with the state of affairs (not necessarily a state of X) which it is intended to convey to Y; X's performance, that is, would be suggestive, in some recognizable way, of the state of affairs without being a natural response of X to that state of affairs.

We reach, then, a stage in which the communication vehicles do not have to be, initially, natural signs of that which they are used to communicate; provided a bit of behavior could be expected to be seen by the receiving creature as having a discernible connection with a

particular piece of information, then that bit of behavior will be usable by the transmitting creature, provided that that creature can place a fair bet on the connection being made by the receiving creature. Any link will do, provided it is detectable by the receiver, and the looser the links creatures are in a position to use, the greater the freedom they will have as communicators, since they will be less and less restricted by the need to rely on prior natural connections. The widest possible range is given where creatures use for these purposes a range of communication devices which have no antecedent connections at all with the things that they communicate or represent, and the connection is simply made because the knowledge, or supposition, or assumption, of such an artificial connection is prearranged and foreknown. Here creatures can simply cash in, as it were, on the stock of semantic information which has already been built into them at some previous stage.

In some cases, the artificial communication devices might have certain other features too, over and above the one of being artificial: they might, for example, involve a finite number of fundamental, focal, elementary, root devices and a finite set of modes or forms of combination (combinatory operations, if you like) which are capable of being used over and over again. In these cases, the creatures will have, or be near to having, what some people have thought to be characteristic of a language: namely, a communication system with a finite set of initial devices, together with semantic provisions for them, and a finite set of different syntactical operations or combinations, and an understanding of what the functions of those modes of combination are. As a result, they can generate an infinite set of sentences or complex communication devices, together with a correspondingly infinite set of things to be communicated, as it were.

So, by proceeding in this teleological kind of way, we seem to have provided some rationale both for the kind of characterization of speaker's meaning which I went in for long ago, and also for the characterization of various kinds of communication systems, culminating in things which have features which are ordinarily supposed (more or less correctly, I would imagine) to be the features of a fully developed language. I say that we *seem* to have provided a rationale; for there is, I think, a large residual question of a methodological kind. My succession of stages is not, of course, intended to be a historical or genetic account of the development of communication and language; it is a myth designed, among other things, to exhibit the conceptual

link between natural and nonnatural meaning. But how can such a link be explained by a *myth*? This question is perhaps paralleled, as was recently suggested to me, by the question how the nature and validity of political obligation (or perhaps even of moral obligation) can possibly be explained by a *mythical* social contract. While the parallel may be suggestive and useful, one might be pardoned for wondering how much more it does besides match one mystery with another. But that is a problem for another day.

### III. The Mystery Package

Well, this is the mystery package. First, a small anecdote. My sometimes mischievous friend Richard Grandy once said, in connection with some other occasion on which I was talking, that to represent my remarks, it would be necessary to introduce a new form of speech act, or a new operator, which was to be called the operator of *quessertion*. It is to be read as "It is perhaps possible that someone might assert that . . .", and is to be symbolized "?-"; possibly it might even be iterable. I treasure this suggestion to just about the same degree as I treasure his dictum, delivered on another occasion, that I "can always be relied upon to rally to the defense of an 'under-dogma.'" Everything I shall suggest here is highly quessertable. I shall simply explore an idea; I do not know whether I want to subscribe to it or not. In what follows, then, I am not to be taken as making any ground-floor assertions at all: except for the assertion that something is quesserted.

The general idea that I want to explore, and which seems to me to have some plausibility, is that something has been left out, by me and perhaps by others too, in the analyses, definitions, expansions, and so on, of semantic notions, and particularly various notions of meaning. What has been left out has in fact been left out because it is something which everyone regards with horror, at least when in a scientific or theoretical frame of mind: the notion of value.

Though I think that in general we want to keep value notions out of our philosophical and scientific inquiries (and some would say out of everything else), we might consider what would happen if we relaxed this prohibition to some extent. If we did, there is a whole range of different kinds of value predicates or expressions which might be admitted in different types of case. To avoid having to choose between them, I am just going to use as a predicate the word "optimal": the

meaning of which could of course be more precisely characterized later.

The reason why I am particularly interested in this general idea is that my own position, which I am not going to try to state or defend in any detail at the moment, is that the notion of value is absolutely crucial to the idea of rationality, or of a rational being. There are many ways in which one can characterize what it is to be a rational being. Some of them may turn out to be equivalent to one another in some sense: they may turn out to apply to exactly the same cases. However, it may be that even though they are equivalent, there is one that is particularly fruitful from a deductive-theoretical point of view. I have strong suspicions that the most fruitful idea is the idea that a rational creature is a creature which evaluates, and that the other possible characterizations may turn out to be co-extensive with this, though in some sense less leading. I do not know whether it follows from this, but at any rate I think it is true, that all naturalistic attempts at the characterization of rationality are doomed to failure. Value is in there from the beginning, and one cannot get it out. This is not something to be argued about here, but it should give some indication of the kind of murky framework in which I shall now start to operate.

It seems to me that there are two different problems connected with meaning in which questions of value might arise. I call them the minor problem and the major problem. The minor problem has to do with the relation between what, speaking generally, I may call word meaning and speaker meaning. It seems plausible to suppose that to say that a sentence (word, expression) means something (to say that "John is a bachelor" means that John is an unmarried male, or whatever it is) is to be somehow understood in terms of what particular users of that sentence (word, expression) mean on particular occasions. The first possible construal of this is rather crude: namely, that usually people do use this sentence, etc., in this way. A construal which seems to me rather better is that it is conventional to use this sentence in this way; and there are many others.

Now I do not think that even the most subtle or sophisticated interpretation of this construal will do, because I do not think that meaning is essentially connected with convention. What it is essentially connected with is some way of fixing what sentences mean: convention is indeed one of these ways, but it is not the only one. I can invent a language, call it Deutero-Esperanto, which nobody ever

speaks. That makes me the authority, and I can lay down what is proper. Notice that we are immediately arriving at some form of evaluative notion: namely, what it is proper to do.

The general suggestion would therefore be that to say what a word means in a language is to say what it is in general optimal for speakers of that language to do with that word, or what use they are to make of it; what particular intentions on particular occasions it is proper for them to have, or optimal for them to have. Of course, there is no suggestion that they *always* have to have those intentions: it would merely be optimal, *ceteris paribus*, for them to have them. As regards what is optimal in any particular kind of case, there would have to be a cash value, an account of why this is optimal. There might be a whole range of different accounts. For example, it might be that it is conventional to use this word in this way; it might be that it is conventional among some privileged class to use it in this way—what some technical term in biology means is not a matter for the general public but for biologists; it might be, when an invented language is involved, that it is what is laid down by its inventor. However, what we get in every case, as a unification of all these accounts, is the optimality or propriety of a certain form of behavior. That concludes my discussion of the minor problem.

The major problem in which questions of value arise has to do not with attempts to exhibit the relation between word meaning and speaker meaning, but with attempts to exhibit the anatomy of speaker meaning itself. At this point, my general strategy was to look for the kind of regresses which Schiffer and others have claimed to detect concealed beneath the glossy surface of my writings on meaning: infinite and vicious regresses which they propose to cast out, substituting another regressive notion, such as mutual knowledge, instead; raising somewhat the question why their regresses are good regresses and mine are bad ones.

However, as I tried, by looking at Schiffer, to disentangle exactly what the alleged regresses are, I found it almost impossible to do so. That is, someone who alleges an infinite regress ought surely to provide a general method for generating the next stage of the regress out of the previous one, and I could see no general way of doing this: the connections between one stage and another seemed to be disparate. That is not to say that there is no way of doing this. I used to think when talking to Schiffer that there was one, and that I understood it, but now I do not. However, since the actual nature of the regress, or

regressive accusation, does not matter very much, what I have done is to invent my own, which I will call a pseudo-Schiffer regress; and so far from trying to make it leaky or creaky at the joints, on the contrary, I would like it to be as strong as I can make it, and if it is not strong enough, I am going to pretend that it is.

The regress can be reconstructed along the following lines. One might start with the idea that when some speaker S utters some sentence to a hearer A, meaning by it that p, he does this wanting A to think 'p'. That is, at stage one we have "S wants A to think 'p,'" where p represents the content of A's thought or intended thought. However, for reasons that came up long ago, having to do with the distinction between natural and nonnatural meaning, we cannot stop at stage one. We have to proceed to stage two, at which we get "S wants A to think 'p, on the strength of the idea that S wants A to think 'p'" ; and so on.

We have now reached a curious situation, in that there is a certain sort of disparity between what S wants A to think: namely "p, on the strength of S wanting A to think that p," and the accounts that are given, so to speak, in the subclause of this as to what it is that S wants A to think. That is, S wants A to think not just "p," but "p on the grounds that S wants A to think that p"; but when we are stating what the grounds are, what A is supposed to think, we find that he is only supposed to think "p." In other words, what we specify A as intended to think of as the reason for thinking "p" is always one stage behind what the speaker envisages as the reason why he wants A to think "p."

We thus arrive at something of the form "S wants A to think 'p, because S wants A to think 'p, because S wants . . .'" and so on. We put in the extra clause in order to catch up, but we never do catch up, because by putting in the extra clause we merely introduce another thing to catch up with. It is like moving from stage one to stage two: we start with stage one, we add the move from stage one to stage two, but by the time we get to stage two, the place we have to get to is stage three, and so on.

I have chosen this regress because it is rather colorful, but it is not the only one I could have used. What I am looking for is an infinite regress which combines the two following characteristics: first, like all infinite regresses, it cannot be realized: that is, a completion, a situation in which S has an intention which is infinitely expanded in this kind of way, cannot actually exist; and second, the idea that it

should exist is a desideratum. That is, what I am looking for is a situation in which a certain highly complex intention is at one and the same time logically impossible and also desirable. That is not in itself, it seems to me, an unreasonable goal: it certainly can happen that things are logically impossible but desirable, and if it does, I can make use of it.

The pattern of analysis which I would now suggest as the primary interpretant for speaker's meaning would be that S is in that state with respect to whatever he wishes to communicate or impart (p) which is optimal for somebody communicating p. It then turns out that when you cash the value of what it is that is optimal, you find that the optimal state is a state that is in fact logically impossible. That in itself seems to me to be not in the least objectionable so far, although there are some points that would have to be argued for. The whole idea of using expressions which are explained in terms of ideal limits would seem to me to operate in this way. The ideal limits might not be realizable in any domain, or they might be realizable in certain domains but not in the domain under consideration; for instance, the fact that they are not realizable might be contingent, or it might be non-contingent. It might be for one reason or another (let us pretend for simplicity) that there cannot be in the sublunary world any things that are, strictly speaking, circular. Nevertheless, that does not prevent us from applying the word "circular" in the sublunary world, because we apply it in virtue of approach to, or approximation to, the ideal limit which is itself not realized. All we need is a way of so to speak measuring up actual particulars against the unrealizable quality of the perfect particular. Indeed, maybe something like this is what Plato went in for.

It seems to me that the notion of knowledge might be explicable in this way. This is a notion which might be, is conceivably, realizable in a certain domain but not in others. Here we look to the people in the past who have suggested that the standard, or crucial, feature of knowledge is that if you know something, you cannot be wrong. Some people then went on to say that it follows that the only things we can know are necessary truths, because there, in some sense of "cannot", we cannot be wrong; and there are various familiar objections to this. Now I might want to say that those people are right, if what they meant was that *strictly speaking* the only things that can be known are necessary truths. However, that does not restrict us to supposing that people who talk about knowing other things are using

the word "know" improperly: all it requires is that there should be some license to apply the word nonstrictly to things which in some way approach or approximate to the ideal cases.

It is not my business on this occasion to suggest exactly, or in any detail, what the demands for approach or approximation might be. I will only say that, whatever they are, they ought to be ones which justify us in *deeming* certain cases to satisfy a given ideal even though they do not, in fact, strictly speaking exemplify it; just as in Oxford on one occasion, there was a difficulty between an incoming provost and a college rule that dogs were not allowed in college: the governing body passed a resolution deeming the new provost's dog to be a cat. I suspect that crucially we do a lot of deeming, though perhaps not always in such an entertaining fashion.

Let me summarize the position we have now reached. First, on this account of speaker meaning, as a first approximation to what we mean by saying that a speaker, by something he says, on a particular occasion, means that *p*, is that he is in the optimal state with respect to communicating, or if you like, to communicating that *p*. Second, that the optimal state, the state in which he has an infinite set of intentions, is in principle unrealizable, so that he does not *strictly speaking* mean that *p*. However, he is in a situation which is such that it is legitimate, or perhaps even mandatory, for us to deem him to satisfy this unfulfillable condition.

Finally, there is the question of how this relates to the regresses which people have actually found: regresses, or prolongations of the set of conditions, which actually exist. Certain ingenious people, such as Strawson and Dennis Stampe, and ending up with Schiffer, who moves so fast and intricately that one can hardly keep up with him, have produced counterexamples to my original interpretant in an analysis of meaning, counterexamples which are supposed to show that my conditions, or any expansion of them, are insufficient to provide an account of speaker meaning. The alleged counterexample is always such that it satisfies the conditions on speaker meaning as set forward so far, but that the speaker is nevertheless supposed to have what I might call a sneaky intention. That is, in the first and most obvious case, his intention is that the hearer should in fact accept *p* on such and such grounds, but should *think* that he is supposed to accept *p* not on those grounds but on some other grounds. That is, the hearer is represented, at some level or other of embedding, as having, or being intended to have, or being intended to think himself

intended to have (or . . . ), a misapprehension with regard to what is expected of him. He thinks he is supposed to proceed in one way, whereas really he is supposed to proceed in another. I would then want to say that the effect of the appearance of a sneaky intention, the function that such a sneaky intention would have in the scheme I am suggesting, would simply be to cancel the license to deem what the speaker is doing to be a case of meaning on this particular occasion: that is, to cancel the idea that this is to be allowed to count as a sublunary performance, so to speak, of the infinite set of intentions which is only celestially realizable.

In a way, what this suggestion does, or would do if it were otherwise acceptable, is to confer a rationale upon a proposal which I actually did make in an earlier paper, to the effect that what was really required in a full account of speaker meaning was the *absence* of a certain kind of intention. This may very well be right, but the deficiency in that proposal was that it gave no explanation of *why* this was a reasonable condition to put into an account of speaker meaning. I think, if we accepted the framework I have just outlined, this arbitrariness, or *ad hocness*, would be removed, or at least mitigated.