

## UNIT V : INTRODUCING GRICE

In the previous unit, I presented two theses which taken together give the skeletal outline of a neo-Lockean theory of meaning. The first, which incorporates the basic contention of agent semantics was this:

- (A) The linguistic type X current in speech community C means that p iff members of community C standardly use tokens of the type X to communicate the thought that p.

The second thesis proffered an analysis of the pivotal concept of "communicating a thought":

- (C) Speaker S uses token x to communicate the thought that p iff S utters x intending thereby to produce in his audience the belief that p (or perhaps: intending thereby to produce in his audience at least the belief that he, S, believes that p).

The aim of this unit is to begin the task of refining these initial crude versions into a more defensible neo-Lockean pair. Our inspiration will be the work of Paul Grice, and in particular the seminal paper 'Meaning' (first published 1957). As will soon become only too apparent, Grice's original paper has prompted an enormous secondary literature, and we will have to spend a lot of time in subsequent units considering various proposed revisions, extensions and objections to Grice's work. This unit, however, will remain pretty close to being basic exposition of Grice's original position. In so far as we do consider a number of initial objections, this will be in order to sharpen up our appreciation of the content and force of Grice's position, and to point to some developments that are perhaps needed to augment the original position but do not supersede it.

Before introducing Grice, four comments. Firstly, it is absolutely crucial to keep separate the question of the viability of (A), or some refined replacement, from the question of the soundness of (C), or some improved version of (C).

The issue about (A) is whether we can give an account of the way linguistic expressions have meaning in terms of what speakers do with those expressions, and if so whether the crucial thing speakers do with linguistic tokens is use them for communication. The issue about (C) is whether it, or its more refined successors, give an adequate analysis of the concept of communication. One could accept, for example, the Gricean revision of (C), while rejecting the attempt to use it in a programme of agent semantics. Conversely, one could welcome some successor of (A) while rejecting Grice's version of (C). So we must try to keep as separate as possible our evaluations of the programme of agent semantics and the quite distinct Gricean analysis of communication.

Secondly, we will see in a moment that Grice in his versions of (A) and (C) does not use the notion of "communicating the thought that p" as the linking concept. Rather he speaks simply of an agent "meaning that p" by what he utters. He presents therefore an analysis of communication as an analysis of what is for a speaker to mean something by an utterance. The snag with this, it seems to me, is that it is far from clear that we have a pre-philosophical notion of a speaker's meaning to be analysed in this way. Of course, we do ask from time to time what a speaker meant by a certain utterance: but here we are standardly enquiring about the insinuations or hidden implications a speaker wants to convey, not about the direct content he attaches to the words. Yet it is the notion of directly attaching a content (of conveying a thought) that Grice is intending to illuminate. It would be to hang a weighty thesis on a slender thread if we made anything turn on the use of a perhaps suspicious locution such as "a speaker's meaning something by an utterance". Hence my preferred strategy of avoiding the locution and talking directly of communication.

Thirdly, to re-emphasize the point of the first section of Unit IV: the interest of (A) and (C), or their descendant variants, does not have to lie in reductionist ambitions. In fact, I reject Bennett's full-scale reductionism for a stratified programme where, at stage (a), we give a basic account of the nature of intentional states of a simple kind, then go on to stage (b), the explication of the additional complexities that set in when language comes onto the scene. But even if this is a mistake, and we cannot give a coherent account of intentional states that can be peeled off from the account of language use, still the interconnections between meaning and psychological states will be of great interest, and getting clear about them a necessary first step in the analysis meaning.

Fourthly, let me briefly introduce a bit of Gricean jargon, the natural/non-natural distinction. Grice begins his 1957 paper by distinguishing two senses of the word 'means' and its cognates, viz. (1) the natural sense, which is the sense the word has in contexts such as "Those clouds mean that it will rain" or "These spots mean he's got measles" and (2) the non-natural sense which the word has in contexts such as "Those three rings on the bell (of the bus) mean that the bus is full" or "That sign means 'Road works ahead'". The distinction noted here seems intuitively sound; it seems fairly clear that there is a difference between (a) taking the occurrence of one thing to be a (natural) sign or indication of the occurrence of something else, and (b) taking something to be a communication device. This distinction shows up in the divergent entailments of the two uses of 'means'. Thus, for example, the truth of "Those clouds mean that it will rain" entails that it will rain, while "Those three rings mean that the bus is full" would remain true even if the bus were not full. By drawing a number of such general contrasts between

the two kinds of use of the word 'mean', Grice is able to confirm - in advance of any deep analysis - that there is indeed a genuine distinction to be made here. We will occasionally follow Grice by marking occurrences of 'mean' with the suffix '-nn' if the word carries (something like) its non-natural sense. For more on the natural/non-natural distinction, see Stampe's fine paper 'Toward a Grammar of Meaning' Phil. Rev. 1968.

(1) Expanding the (A) - thesis

As we have just seen, the word 'mean' can occur in either its natural or its non-natural sense in sentences of the form "O means ..." where 'O' stands for some object. But in the context "x meant ... (on a particular utterance occasion)" where x is an uttered token, 'meant' will normally have its non-natural sense. And Grice suggests the following equivalence

- (M) "x meant-nn so-and-so (on a particular utterance occasion)" is (roughly) equivalent to "Somebody meant-nn so-and-so by x (on that particular utterance occasion)".

Here, on the right of the equivalence, Grice employs his notion (perhaps a dubious one) of somebody's meaning something - where I, perhaps, would prefer to speak of someone's communicating something (or rather, I should say aiming to communicate something: for the verb "communicate" suggests success, but evidently a speaker still attaches significance to his words even if he fails to convey his thought because of incomprehension or inattention on the part of his intended audience). Grice is going to go on to offer a (C)-style account of his pivotal notion of a speaker's meaning, which together with (M) will deliver an account of what it is for an uttered token to have a meaning. How, next, do we move from an account of what it is for a linguistic type to have meaning? In other words, how do we proceed to an account of 'means' as it occurs in the claim 'The Welsh sentence "Mae glo yn ddu" means coal is black'? Grice suggests in effect that (a) "X means-nn (timeless) so-and-so" - where X is an utterance type - is roughly equivalent to (b) "In general, particular tokens x of the type X mean-nn (on their occasion of utterance) so-and-so." And this together with (M) and (G) yields an analysis of (a) as a complex statement about a group of utterers' intentions: as Grice puts it

- (N) "X means-nn (timeless) that so-and-so" might as a first shot be equated with some statement or disjunction of statements about what people (vague) intend to effect by (tokens of the type) X".

Grice is here sketching the following programme for analysing the concept of linguistic meaning. We are to start with a (C)-type analysis of

- (I) A intended to communicate something to an audience by uttering x (on a particular occasion)

or in Grice's preferred phraseology

- (Ia) A meant-nn something by uttering x (on a particular occasion;

and our initial analysis is not to have recourse to the notion of conventional meaning as used, for example, in reporting the sense of the Welsh sentence 'Mae glo yn ddu'. As Stampe has put it, "the great gain of such an analysis as Grice's is that, whatever its difficulties, the concept of intention (the key ingredient of (C)-type analyses), unlike the concepts of an idea or concept, semantic marker, semantic regularity, and so forth, at least does not swim in the same orbit of conceptual space as does 'meaning' itself". But now, suppose we have found an adequate Gricean analysis of (Ia) - then, Grice implies, we could use this analysis of the notion of A's meaning something by the act of uttering x to elucidate the notion of A's meaning something by the token uttered. This would give us an analytical account of

(II) A meant something by x (on a particular utterance occasion)

If we add to this a recipe for determining what A meant given his particular communication intentions, then we will have an account of

(III) A meant so-and-so by x (on a particular utterance occasion)

We could then use this result to elucidate the notion of someone's usually meaning so-and-so by tokens of the type X, which would give us an account of

(IV) A usually means so-and-so by X

This in turn can be used to elucidate

(V) Members of community C usually mean so-and-so by X

And finally, Grice implies, our account of (V) can be used to elucidate the notion of a sentence's conventionally meaning something in community C - that is, to elucidate

(VI) X means so-and-so (in the language of community C)

In short, then, Grice implies that (1) we can analyse (Ia) without relying on a prior understanding of the notion of conventional meaning, and (2) we can then use this analysis to give an illuminating analytical account of (VI) - and hence of the notion of conventional meaning.

Before turning to see how Grice fares with the first stage of improving on the (C)-thesis account of (I) or (Ia), let us look at the couple of objections to the idea that we can illuminate (VI) by means of the type of connections which Grice has sketched that lead back to (I).

## (2) Ziff's Objection

In the chain of moves from (Ia) to (VI), it is the last link which is the most controversial. Is it really acceptable? Can conventional linguistic meaning be reduced in this way to an aggregate of meanings-on-a-particular-occasion?

Well, as Ziff (Analysis 1967) and others have pointed out, Grice's first shot equivalence between (V) and (VI) won't do

at all. For we sometimes can say that a sentence X means p (in a given language L) even though we do not believe that if an L-user were to utter X he would thereby usually mean that p. Consider first an example of Bennett's. In English, the sentence "No head-injury is too trivial to ignore" has a certain meaning; yet empirical investigation shows that most English-speakers, if they uttered this sentence, or read it, or heard it uttered, would give it a quite different meaning that is in fact the meaning of "No head-injury is so trivial that it can be ignored". (To see that these two sentences do not mean the same, cf. "No brandy is too old to drink" and "No brandy is so old that it can be drunk"). Another example: the sentence "He is the son of a stickleback fish" has a certain meaning - it says that the male in question is the offspring of a small scaleless fish having two or more free spines in front of the dorsal fin etc. But it is not at all likely that an English-speaker uttering the sentence would mean that by it. As Ziff (whose example this is) puts it: "most likely, by uttering this sentence people would intend to denigrate a contextually indicated male person". So, we clearly do distinguish in particular cases between what a sentence means in a given language and what most members of that language community mean by it, or would mean by it if they uttered it. Grice's original equivalence therefore needs to be complicated in order to deal adequately with this distinction.

However, I suggest that the non-equivalence between (V) and (VI) can be explained and justified within a Gricean framework. To see this, a few reflections about word-meaning are in order (and here I follow closely an argument strategy used by Bennett in lectures in 1966).

Consider again the claims about the relationship between sentence-meaning and word-meaning which I made in I, pp.15-17. There maintained that in an important sense sentence-meaning is primitive, and words have meaning just in so far as they make a systematic contribution to the sense of sentences containing them. Clearly, this position is not merely compatible with, but a necessary adjunct to, Grice's programme.

If we are attempting to elucidate the notion of meaning through the notion of an utterer's communication-intentions, then clearly we will have to treat sentence-meaning as in a sense primitive. For it is the complete utterance which is the subject of the communication-intention, the means by which the utterer is attempting to effect his intention. Thus, if communication-intentions are the starting point of our analysis of meaning, then it is the notion of the meaning of complete utterances which is secured first in our programme of analysis - the notion of word-meaning being developed from this.

I sketched, at I p.17, an outline account of how a child might come to grasp the meaning of, say, the word 'milk' via a grasp of sentences involving that word. I stick to that story - and a similar story can likewise be told to deal with

the child's grasp of the significance of syntactic features of utterances such as word-order. And what goes for the child goes for the linguist-in-the-field engaged in the "radical translation" of a previously unencountered language. The linguist proceeds from a provisional, fragmentary, and no doubt vague 'sentence-dictionary' for the native language, via the discovery of correlations between aspects of native sentences and aspects of their supposed communication content, to a provisional grammar and word-dictionary for the language. This can be tested against further native utterances, and so further refined, etc. However, telling the story in terms of the child learning its own language or the linguist engaged in "radical translation" is perhaps misleading. For after all, it is a contingent fact that we have to learn our language as children; and it is another contingent fact that there happens to be more than one human language for the linguist to practise on. Indeed as Armstrong reminds us, it is conceivable that both should have been false, and "whole of syntax and semantics should have been innate so that mankind naturally spoke the one, wired-in language". So, rather than telling the story in terms of some-one learning a language, we should perhaps re-tell it in terms of a speaker's method for checking his correct grasp of another's speech. Still, the former way of presenting the story lends clarity and drama to a complex tale.

Staying with the radical translator for a while, we can envisage the possibility that his most satisfactory grammar-plus-word-dictionary, while accommodating most of his data, still leaves some data uncovered. This might give him grounds for saying that the native use of the particular sentences involved in the unaccounted data was deviant - in particular, the translator might have a ground for distinguishing 'what the sentence S means in the native language' from 'what the natives usually use the sentence S to mean'. And this possible situation provides us, I think, with a model of how we arrive at a distinction which we do in practice make occasionally between the 'real' meaning of, say, an English sentence, and what English speakers typically use the sentence to mean.

Let us spell this possibility out in a Gricean way. Consider a community whose members go in for uttering 'sentences' intending thereby to communicate with one another. As we saw, it is Grice's basic contention that we can give an account of the notion of meaning (to communicate) something by an utterance without bringing in at this stage the notion of conventional meaning. Now, suppose that there are observable regularities in the communication situation between sentence uttered and intended effect, and suppose also that the community in general censures those communication performances which depart from the normal regularities. Then we may be able to bring our data concerning the 'linguistic' performances in the community under a general theory which says that these data are uses of language in accordance with a set of 'meaning'-rules R - i.e. a set of rules assigning to sentence-forms the communication intentions they are to be used to effect. Of course, in saying that the data are uses of language in accordance with such-and-such rules, it is not being suggested that the performers are consciously following the rules. Their 'having' the rules may show up only in the observed regularities in performance and their criticisms

of variations from the norm (based on the performers mutual expectations with regard to the behaviour of others). Now it may happen that we find (1) that a very large portion of our data is perfectly explained by the theory that the performers of our community have certain meaning-rules R and stick by them, but (2) that some further data does not conform to this theory. This may of course simply be evidence that we are wrong about what the community's linguistic rules are. But we are not forced to that conclusion if we can unarbitrarily explain the deviant linguistic data in other ways. Such explanations are the province of a 'theory of performance', in Chomsky's phrase. They may refer, for example, to the language-users' difficulty in understanding (and resistance to using) expressions which go beyond certain limits of length and complexity, or to the language-users' reluctance to construe literally utterances which are quite ridiculous when so understood, and so on. It would be absurd for us to insist that all our data must be accommodated by our theory of meaning for the community's utterances if we can reasonably sift out and discard some of the data by appealing to our theory of performance. Consider again our two examples from English: Firstly, it is empirically the case that the sentence (a) "No head-injury is too trivial to ignore" is generally taken to be equivalent to (b) "No head-injury is so trivial that it can be ignored". Are we to explain this by taking (a) to be a special idiom, or by claiming that some word does not have its standard meaning in (a)? Surely not, for we may give a reasonable explanation of the deviant use of (a) along the following lines: the sentence (a) has a certain complexity akin to the complexity of sentences containing multiple negations and so is rather apt to be misunderstood; this combined with the fact that the sentence says something blatantly false when construed literally explains our tendency to misconstrue it as expressing the much more reasonable proposition which is properly expressed by the similar sentence (b). Secondly, Ziff's example: Surely the fact that we would naturally interpret (c) "He is the son of a stickleback fish" as expressing denigration - at least when the pronoun refers to a person and not to a fish - is explained by the facts the (c) would say something quite ridiculous if interpreted literally and that (c), while not itself an idiom, is closely akin to various idiomatic expressions which are used to denigrate people. So, in these two cases, we may appeal to considerations which constitute a partial theory of performance in order to deal with data which do not fit easily into a semantics for English. And in general, given a language-community who usually use a sentence S in a way which is not in accordance with our theory of meaning for their language L, it may be possible to handle this deviation through a theory of performance rather than treat it as a genuine counter-example to our theory of meaning. And if this is possible, then we will have adequate grounds for distinguishing 'What S means' - i.e. its meaning as given by the meaning rules for L from 'what people (i.e. L-users) usually mean by S'.

So, Ziff is correct to object to the over-simple equation of (V) with (VI). But now note that this does not undermine Grice's programmatic insight that the notion of 'what X means' must be elucidated via the notion of 'what people generally mean by X'. I have sketched above a way in which we might come

to distinguish 'what X means' from 'what people usually mean by X' while remaining within a Gricean framework. On this account, our theory of what sentences in a language L mean is still based on facts about what L-users usually mean by their various utterances - to quote Bennett (in lectures):

*The theory of performance comes in simply as a way of disciplining the selection we make from the totality of available facts about the linguistic behaviour of L-users ... the final court of appeal in determining what a sentence X means in L must be what L-users mean by some sentences of L if not what they mean by X itself.*

We must therefore reject the crude equation between (V) and (VI): but this is in fact compatible with continued support of Grice's programmatic suggestion that an account of (V) can be used as the basis of an account of what it is for a sentence-type to have a certain standard meaning.

### 3. Quinton's objection.

Consider even the unrevised (C) thesis: according to this an utterer A (aims to) communicate the thought that p only if he intends to produce in his audience R a certain response r, which is in fact the acquisition of a certain belief. However, one generally cannot intend to produce some result by a given means, if one does not believe that there is any likelihood that the means will be effective in securing the desired result. (Consider the case of someone pushing against the side of a house who claims (a) that he is intending to push the house down, but (b) doesn't believe that it is possible to demolish the house by pushing against it). So, it seems that a necessary condition of A's aiming to communicate that p (or, as Grice would put it, meaning-nn that p), is that A should believe that there is some likelihood that his uttering x will produce the desired response r. But what grounds could A possibly have for this belief? Leaving aside the small but important class of iconic utterances (where, for instance, A mimes being stuck on the head by a falling coconut to warn another of the impending danger of falling coconuts), it seems that A's ground for believing that uttering x will secure response r will simply be the belief that R will recognize what the speaker is up to, which surely requires R to know the conventional meaning of A's utterance.

Let me expand on that. Perhaps in the iconic case, A can hope that his audience recognizes what he is up to in virtue of the imitative or pictorial character of A's utterance (though it should be pointed out that even here, A is perhaps unwittingly relying on the assumption that he and his audience will generally share the complex disposition to 'understand' a picture in the same way - cf. Investigations II.xi). But in the case of most utterances of sentences of a natural language, there is no such pictorial element. So what possible reason could A have for thinking that his uttering a given sequence of sounds (say, a sequence that constitutes a token of the English sentence-type 'You are in danger of being hit on the head by falling coconuts') should

stand some chance of having the remarkably complex effect of getting his audience to adopt a particular belief (say, one to do with the danger of falling coconuts)? If A isn't to be seen as making a wild shot in the dark, it seems that his confidence that uttering certain sounds will successfully evoke just such-and-such effect in R must be based on a belief that standardly R just does respond in the required way to just those sounds. And if A's confidence extends even to the case of sentences which he has not tried out on R before, then it seems that it must be based on some general attribution to R of knowledge of the standard use of, say, English sentences for communication purposes.

In short, then, a consideration of how A can reasonably have the intentions attributed to him by the analysis (C) - or its more complex variants - soon forces us to introduce into the picture standard or conventional means for effecting communication intentions. For in most cases, A can only have the intentions needed to satisfy (C) if he does avail himself of means he can reasonably rely on to effect his intentions, and this will usually involve availing himself of the resources of a natural language he shares with his audience.

However, it might be objected that we are going round in circles here. For Grice is offering an analysis of 'A meant something by x' as a basis for a programme whose aim is to elucidate the notion of conventional meaning. Yet I am now saying that in many cases, A cannot reasonably have the intentions required of him by (C) or the more sophisticated Gricean variants unless he avails himself of the resources of a natural language whose sentences have, of course, a conventional meaning. Doesn't this render Grice's whole programme circular?

Critics have certainly maintained that the programme does come to grief at this point, floundering upon a vicious circularity. Here's Anthony Quinton:

*I am convinced that Grice's project is misconceived. The reason is that for someone to intend, by doing A, to bring about B, he must believe that doing A is likely to bring about B. I cannot say that I am intending to make the bus go faster while it is stuck at the traffic lights by tapping my kneecap unless I believe that doing so is calculated to have that effect. Now how can I intend to bring about the belief in you that it is raining by uttering to you the sentence 'it is raining'? Only if I believe or take 'it is raining' to mean what it does. Meaning cannot be defined in terms of Gricean intentions since the possession of meaning by the verbal instruments involved is necessarily presupposed by making use of them for the carrying out of the intentions in question.*

But is this accusation of circularity in fact sound? Must the notion of conventional meaning be surreptitiously involved in the analysis of speaker's occasion meaning in such a way as to make Grice's project entirely misconceived?

The answer is 'No'. Grice's basic contention is (to quote from Strawson's sympathetic exposition in 'Meaning and Truth') that

*... it is possible to expound a concept of communication-intention or, as he calls it, utterer's meaning, which is proof against objection and which does not presuppose the notion of linguistic meaning.*

And this Grice has given good reason to think is possible. Grice's refined versions of (C) are progressively more plausible, and certainly do not explicitly presuppose the notion of conventional linguistic meaning. The latter notion comes into the picture, according to Grice, essentially when we have a community whose members habitually use certain utterances to effect certain communication-intentions. Given that a certain habit is partially established - say to use utterances of the type Y to convey that p - then an utterer has some reason to think that another utterance of Y will successfully convey that p, and hence some reason to use Y when he wants to convey that p. And given the premium on successful communication, the community will have some reason to criticize members who break from the practice of using Y-utterances to convey that p. In short, a past habit will tend to persist and become further entrenched. Thus, the practice of using Y-utterances to convey that p will tend to remain alive because of the interlocking mutual expectations of members of the community - an audience will expect an utterer to use Y-utterances to convey that p, an utterer will expect his audience to have that expectation, and so on. There being such interlocking mutual expectations giving continued life to a practice is just what is necessary (though not sufficient) for the practice to be not merely a regular habit but a convention. This thesis is argued for with great skill by David Lewis in his magnificent book "Convention". Here, Lewis provides a welcome liberation from the sub-Wittgensteinian mystification that has surrounded such notions as 'convention' and 'rule' in much recent philosophical writing. But getting back to Grice, he can allow that over the passage of time, habitual communication devices (which we may suppose were initially iconic, or substitutions for iconic devices) can become more and more complex until we reach the point where an utterer can form communication-intentions which he cannot reasonably hope to effect except by using the conventionally established device for so doing. And in the case of a speaker of a natural language, most of his communication-intentions will be of this kind, as was pointed out above. But this concession should now be seen not to undermine the claims (1) that the fundamental notion of utterer's meaning can be elucidated without reference to the notion of conventional meaning, and (2) that the notion of conventional meaning can be understood by seeing a language as a set of conventional devices for effecting communication-intentions (i.e. roughly as a set of devices habitually used by a community to effect communication-intentions).

To get this clearer, I will quote at length Strawson's succinct discussion of these points in "Meaning and Truth":

It is clear that we can, and do, communicate very complicated things by the use of language; and if we are to think of language as, fundamentally, a system of rules for facilitating the achievement of our communication-intentions, and if the analysis is not to be circular, must we not credit ourselves with extremely complicated communication-intentions (or at least desires) independently of having at our disposal the linguistic means of fulfilling those desires? And is this not absurd? I think this is absurd. But the programme of analysis does not require it. All the analysis requires is that we can explain the notion of conventions of communication in terms of the notion of pre-conventional communication at a rather basic level. Given that we can do this, then there is more than one way in which we can start pulling ourselves up by our own linguistic boot-straps ... We can, for example, tell ourselves a story of the analytic-genetic variety. Suppose an utterer achieves a pre-conventional communication success with a given audience by means of an utterance, say x. He has a complex intentions vis-a-vis the audience of the sort which counts as a communication-intention and succeeds in fulfilling that intention by uttering x. Let us suppose that the primary intention was such that the utterer meant that p by uttering x; and since, by hypothesis, he achieved a communication success, he was so understood by his audience. Now if the same communication problem presents itself later to the same utterer in relation to the same audience, the fact, known to both of them, that the utterer meant that p by uttering x before, gives the utterer a reason for uttering x again and the audience a reason for interpreting the utterance in the same way as before. (The reason which each has is the knowledge that the other has the knowledge which he has). So it is easy to see how the utterance of x could become established as between this utterer and this audience as a means of meaning that p. Because it has worked, it becomes established; and then it works because it is established. And it is easy to see how this story could be told so as to involve not just a group of two, but a wider group. So we can have a movement from an utterer pre-conventionally meaning that p by an utterance x to the utterance-type x conventionally meaning that p within a group and thence back to utterer-members of the group meaning that p by a token of the type, but now in accordance with the conventions.

The story whose beginning Strawson sketches can, I think, be convincingly elaborated. But in any case, we have already said enough to block Quinton's accusation of circularity. It is true (as Quinton says) that in complex situations, rationally possessing Gricean communication intentions requires the expectation that the conventional sense of what one utters will be grasped. But, in Grice's view, this is to say no more than that one's utterance is produced at least in part because it is an element in an entrenched communication practice. And the notion of communication, and hence of an entrenched communication practice can be elucidated in Gricean ways, thus giving us an account of conventional meaning.

4. Modifying the (C)-thesis.

People frequently perform actions (such as waving an arm, drawing a diagram, uttering a sentence) with the intention of thereby communicating something to someone else. Purely to simplify the discussion at this stage, let us restrict our attention for the moment to cases where the performer A intends to convey information to his 'audience' and not, say, a directive: In other words, let us consider first cases where A performs an action thereby intending to tell something to his audience. Now clearly there is a sharp distinction to be drawn between (a) performing an action with the intention of thereby telling somebody something and (b) performing an action with the intention of thereby getting someone to believe something. To take Grice's example: I might leave B's handkerchief near the scene of a murder in order to induce in the detective the belief that B was the murderer. This would certainly be a case of performing an action with the intention of thereby getting someone to believe something; but I think it would be clearly wrong to say that this was a case of communicating or trying to communicate information to someone. Trying to fool someone into thinking that B is the murderer is quite a different business from trying to tell him 'straight out' that B is the murderer; communicating is a much more 'open' affair. So, if I try to tell somebody something, then typically I not only intend (i) to produce in him a certain belief - at least according to Grice - but also intend (ii) that he should recognize what I am up to. In the 'handkerchief' case, however, my evil design would have been frustrated if the detective had recognized that the handkerchief hadn't been dropped accidentally but had been left by me with the intention of inducing in him the belief that B was guilty. We need, therefore, to emend the original (C)-thesis to bring out the requirement of openness in genuine cases of communication as against evidence-faking.

Consider next another example from Grice: Herod presents Salome with the head of St. John the Baptist on a plate. Herod presumably (i) intended thereby to produce in Salome the belief that St. John the Baptist was dead and no doubt (ii) also intended Salome to recognize this first intention - i.e. recognize what he was up to. However, it would seem to be a mistake to say that Herod was therefore telling Salome that St. John was dead: there is a difference between showing someone evidence that p is the case and simply telling him that p is the case. If I show someone evidence that p then, while I may intend in this way to get him to believe that p, I do not rely on his taking my word for it but on his accepting the evidence I put before him as evidence that p. On the other hand, if I merely tell someone that p then I normally intend him simply to take my word for it - that is, I intend to get him to believe that p via his recognizing that this is what I intend. Grice has another extremely neat example to illustrate the difference between 'showing evidence for' and merely 'telling': he asks us to compare the following two cases.

- (1) I show Mr. X a photograph of Mr. Y displaying undue familiarity to Mrs. X,
- (2) I draw a picture of Mr. Y behaving in this manner and show it to Mr. X.

Now it seems intuitively clear that case (2) is very like, and case (1) rather unlike, simply telling X that Y had been unduly familiar with Mrs. X. What then is the relevant difference between the two cases? Grice answers: "Surely that in case (1) Mr. X's recognition of my intention to make him believe that there is something between Mr. Y and Mrs. X is (more or less) irrelevant to the production of this effect by the photograph. Mr. X would be led by the photograph at least to suspect Mrs. X even if instead of showing it to him I had left it in his room by accident; and I (the photograph shower) would not be unaware of this. But it will make a difference to the effect of my picture on Mr. X whether or not he takes me to be intending to inform him (make him believe something) about Mrs. X, and not just doodling or trying to produce a work of art." In other words, the relevant difference between the evidence-showing case (1) and the quasi-telling case (2) is that in the latter case Mr. X's recognition of what I am up to (of my intention in showing him the picture) is intended to play a major role in producing the desired effect. So, on the basis of the consideration of this kind of example, we might conjecture that "A intended to tell his audience something (communicate information) by his performing act a" may be analysed as:

"A performed a intending

- (i) to produce some belief in his audience
- (ii) that his audience recognize intention (i)
- (iii) that his audience fulfil (i) at least partly through recognizing A's intention (i)"

We might put this more succinctly but less clearly thus:

"A intended the performance of a to induce a belief in his audience by means of the recognition of this intention".

This sort of analysis may be generalized in an obvious way to cover communication situations other than ones where it is information that is being conveyed. If I am trying to communicate a directive, for example, then the effect I intend to produce in my audience is not a belief but an action (or at any rate, the intention to perform an action). But telling someone to do something involves more than merely trying to produce action on his part. Compare

- (1) A policeman's trying to stop a thief's get-away car by jumping out into its path,
- (2) A policeman on point-duty trying to stop a car by holding up his hand.

In the first case, the policeman intends to get the thief to stop his car - and no doubt intends the thief to recognise this. However, it isn't the thief's recognition of the policeman's purpose which is intended to produce the desired effect; rather, the policeman is relying on the thief's presumed

reluctance to commit homicide (someone's accidentally stepping out into the path of the car would probably have the same effect). In the second case, however, the driver's recognition of the policeman's intention is precisely what is intended to be effective in getting him to stop. This leads us naturally to the conjecture that if I intend to communicate a directive, then I intend (i) to produce an action or intention to act in my audience, and (analogously to the informative case) also intend (ii) that the first intention be recognized by my audience, and further intend (iii) that the intended effect in my audience be produced at least in part by my audience's recognizing that I intend to produce it. Other sorts of communication situation can be dealt with similarly. We might thus suggest that in general "A intended to communicate something to an audience by his performing act a" is equivalent to "A intended the performance of a to produce some effect in his audience by means of the recognition of this intention" - and specifying the effect in question gives us the content of the communication.

Next a minor point: In many cases, A's performance - his act a - will be roughly characterizable as his producing an object; for example, A may utter a sentence, draw a diagram, produce a sequence of noises by a Morse buzzer, produce puffs of smoke etc. In such cases we might say 'A utters something' (this extended use of the verb 'utter' to cover physical activities other than speaking is now fairly standard). And for these cases our putative analysis (in its brief form) will read: "A intended to communicate something to an audience by uttering x" is equivalent to "A intend the utterance of x to produce some effect in an audience by means of the recognition of this intention".

So far, then, we have traced the outline of an account of communication that can be extracted from Grice's 1957 paper. But of course, that paper is entitled 'Meaning', not 'Communication': the presumption here is that an account of communication is an account of meaning. Grice, like Locke before him, sees the communicative uses of language as the clue to an analysis of the notion of linguistic meaning. This presumption is not obviously correct: the notion of a solitary non-communicating language user, whose language is essential to his rationality, is not obviously absurd (for more on this, see my 'Solitary Speakers' Mind 1975). But for the moment we won't question this presumption: even if it is a mistake, the account of communicative uses of language that we will develop in a Gricean spirit will serve as a constraint on our further discussion of meaning. In other words, even if our story about communicative uses of language is not the whole story about language, we can still require our future discussions about language to entail roughly Gricean conclusions about the particular case of communicative uses.

For the moment, then, let's fall in with the Locke/Grice presumption, and bring the notion of meaning into the picture in Grice's way. I have put forward for consideration a Grice-style analysis for

- (1) A intended to communicate something to an audience by uttering x

But surely (I) is at least roughly equivalent to

- (II) A meant something by x (on a particular utterance occasion)

Here we have shifted the focus of attention from the act of uttering to the token uttered but the two statements appear to have (approximately) the same import. But if we accept the Grice-style analysis of (I) and the rough equivalence of (I) with (II), we then have the following:

"A meant something by x' is roughly equivalent to "A intended the utterance of x to produce some effect in an audience by means of the recognition of this intention",

and we might add that to ask what A meant will be to ask essentially for a specification of the intended effect. And this is, of course, precisely Grice's formulation in his 1957 paper. Spelling this out in full we get:

- (G) "A meant something by x (on a particular occasion of utterance)" is (roughly) equivalent to: "For some audience R, A uttered x (on that particular occasion) intending
- (i) to produce in R a particular response r
  - (ii) R to recognize that A intends (i)
  - (iii) R to fulfil (i) - that is, to produce r - at least partly for the reason that he recognizes that A intends (i)".

A specification of r will determine what A meant. For example, if the intended response r is the acquisition of the belief that p, then what A meant by x is (that) p. Let's spell out an example, then:

- (G\*) "A meant (that) Quine is a logician by x (on a particular occasion of utterance)" is (roughly) equivalent to: "For some audience R, A uttered x (on that particular occasion) intending
- (i) to produce in R the belief that Quine is a logician
  - (ii) R to recognize that A intends (i)
  - (iii) R to come to believe that Quine is a logician at least partly for the reason that he recognizes that A intends (i)".

Note that on this account what I may mean on a particular occasion by x may be very different from what (if anything) x itself might ordinarily be said to mean.

## 5. Searle's objections

As we saw in the previous, purely expository section, Grice offers us an analysis of the notion of someone's meaning something by what he says or does, an analysis which doesn't obviously drag in the notion of conventional meaning. In this section, we will begin taking a closer look at the analysis: obviously, if we can't analyse the notion of someone's meaning something by a particular utterance without appeal to the notion of conventional meaning, then Grice's programme is doomed to circularity. In particular, we will begin by examining objections to the sufficiency of Grice's analysis of occasion-meaning.

One such objection is offered by Searle in his paper "What is a Speech Act?" (reprinted in Searle ed. Philosophy of Language) and essentially repeated by him in his book Speech Acts, section 2.6. Here, Searle firstly presents a putative counter-example to (G), and then goes on to suggest an appropriate amendment to (G). As we will see, the amendment he advocates goes clean against the lines of Grice's thought; if we accept Searle's criticisms, then we have to abandon Grice's programme as a lost cause. Fortunately, Searle's case is less than convincing.

First, Searle's argument designed to refute Grice's claim (G) that A means so-and-so by x (on a particular occasion) if and only if A has a certain set of intentions (i), (ii) and (iii): Searle produces a story which is designed to exemplify a case of A's uttering x with intentions (i), (ii) and (iii), but not meaning by x what he would have to mean if (G) were true. Suppose that A is an American soldier in the Second World War who is captured by Italian troops. Hoping that his captors recognize but do not understand German when they hear it, A utters the only sentence of German he knows, namely K "Kennst du das land wo die Zitronen bluen". In this way, A intends to get his captors to believe that he is a German soldier. And according to Searle, A may intend that this response to produce in his captors by their believing that A is trying to tell them that he is a German soldier. In this case, Searle maintains, A intends to produce a certain effect in his audience (namely, their believing him to be a German soldier) by means of the recognition of this intention. So on Grice's account (with the obvious recipe for determining what someone means given the response r he is intending to produce), what A means by K on this occasion is 'I am a German soldier'. But this, Searle maintains, is plain false.

But is this a genuine counter-example of Grice's analysis? Well, for Searle's story to work A must (i) intend to produce in his captors the belief that he is a German soldier, (ii) intend that his captors recognize his intention (i), and (iii) intend that the belief in question is produced in his captors through their recognizing his intention (i). But in what situation would A have this set of intentions? The most natural thing for A to do in his sticky situation is to utter K in an authoritative tone of voice and hope his captors argue somewhat as follows: "He's just said something in German, though we haven't the foggiest idea what. Still, if he speaks German, then most likely he is a German soldier, for what other German would be in this part of the world?" However, as Grice points out in his 1969 paper "Utterer's Meanings and Intentions", Phil. Rev., if A intends his captors to argue like that, then A, while intending the Italians to believe him to be a German soldier, would not be intending them to believe this on the basis of their recognition of his intention. In short, A would not have intention (iii). So, taken in the most natural way, Searle's story is not of someone with intentions (i), (ii) and (iii); indeed, on the natural interpretation, Searle's tale is a straightforward tale about evidence-faking (A tries to get his captors to think he is German by putting on a show of being German).

So, if Searle's story is to be built up into a decent counter-example to (G), then we must avoid the most natural interpretation, and fill in the details in such a way that A does have intention (iii). And this isn't easy. For a start, why should A expect his captors to think that in uttering K he is trying to produce in them the particular belief that he is a German soldier? He could equally well be saying 'Take me to your commanding officer', 'Don't any of you peasants speak German?' etc., etc. Searle has a footnote on p. 44 of Speech Acts which is, I think, designed to rebut this criticism; he suggests that the story may be amplified by adding that A knows or believes that his captors know that there are German soldiers in the area wearing American uniforms, and that his captors have been instructed to be on the look-out for these Germans and to release them as soon as they identify themselves etc. In this case, A can perhaps reasonably expect his captors to have K as meaning that he is a German soldier. But this is only so because he believes that his captors already expect any German sentence spoken by a newly captured soldier in American uniform to mean something like 'I am a German soldier'. So in this case A intends to produce in his audience the belief that he is a German by their recognizing that K is a German sentence. We still haven't got intention (iii) into the picture: indeed, we haven't moved appreciably from the straight evidence-faking case. And it's not easy to see how we can get intention (iii) into the picture. For in general, one can't intend that some result be achieved by means M, if one believes that there is no likelihood that it will be achieved. So, if A is to intend (iii) that the belief that he is a German soldier be produced through his captors recognizing that this is what he intends them to believe, then A must think that there is a likelihood that his captors' recognition that he is trying to get them to believe that he is a German soldier will incline them to believe that he is a German soldier. And in the circumstances, this isn't very reasonable. Why should A think that if it occurs to his captors that he is trying to get them to think that he is German, then this will decrease any inclination they have to think that a deception is being practised upon them?

Let's suppose, however, that Searle can get over this problem, and describe a case where A does have intentions (i), (ii) and (iii). Now it becomes doubtful whether, after all, Searle is right to say that A did not mean 'I am a German soldier'; Grice has a good example:

*The proprietor of a shop full of nickknacks for tourists is standing in his doorway in Port Said, sees a British visitor, and in dulcet tones and with an alluring smile says to him the Arabic for 'You pig of an Englishman'. I should be quite inclined to say that he had meant that the visitor was to come in, or something of the sort. I would not of course be the least inclined to say that he had meant by the words which he uttered that the visitor was to come in - and to point out the German line means not 'I am a German soldier' but 'Knowest thou the land ...' is not relevant. If the Americans could be said to have meant that he was a German soldier, he would have meant that by saying the line, or by saying the line in a particular way just as the Port Said shop-merchant would mean that the visitor was to come in by saying what he said, or by speaking to the visitor in the way he did.*

This seems basically right, though the points made here need a bit of sorting out.

(1) Grice is surely correct about his "Arab merchant" example - the merchant intends by his performance to get the tourist into his shop, and he intends this to come about through the tourist recognizing this intention (and so the merchant satisfies the three conditions of (G) ) - but equally, he can be said to have meant (to communicate) by his performance that the visitor was to come into his shop. The incidental fact that the merchant amuses himself by using in his performance the Arabic for 'You pig of an Englishman' doesn't radically distinguish him from his neighbour who merely gestures his invitations to the tourists. So, to return to the "American soldier" example, Searle's conclusion that it is "plainly false" that A means he is a German soldier (even when he satisfies the three conditions of (G) ) is unobvious and unconvincing. And the reason Searle gives for his conclusion certainly won't do. He says:

*I find myself disinclined to say that when A utters the German sentence what he means is 'I am a German soldier' ... because what the words mean and what A remembers that they mean is 'Knowest thou the land where the lemon trees bloom?'*

Now as it stands, this remark implies that A cannot mean by uttering K something other than what the sentence means in the natural language of which it is a part. And that is surely wrong. Here's Grice:

*I have been listening to a French lesson being given to the small daughter of a friend. I noticed that she thinks that a certain sentence in French means "Help yourself to a piece of cake", though in fact it means something quite different. When there is some cake in the vicinity, I address to her this French sentence and as I intended, she helps herself. I intended her to think (and to think that I intended her to think) that the sentence uttered by me meant 'Help yourself to some cake'; and I would say that the fact that the sentence meant, and was known to me to mean something quite different is no obstacle to my having meant something by my utterance (namely, that she was to have some cake).*

Generalizing this point against Searle, consider

- (a) A utters X meaning thereby that p
- (b) X standardly means that p
- (c) A thinks that X standardly means that p

where X is a sentence in some natural language. Obvious, (a) doesn't imply (b), for A may be mistaken about X's meaning - but Searle's argument (viz. that since A doesn't think that K means 'I am a German soldier', he can't utter K meaning thereby that he is a German soldier) suggests that (a) implies (c). This suggestion is refuted by Grice's example, where only this is true;

- (d) A thinks that his audience thinks that X standardly means that p.

But while (d) is true in Grice's example, it isn't a necessary

condition for the truth of (a). For (a) can be true even when (b), (c) and (d) are false, just so long as

- (e) A thinks that his audience thinks that A thinks X standardly means that p.

(2) To return to Grice's "Arab merchant" case: Grice urges that we should say of the Arab that by uttering his sentence, he meant that the tourist was to enter his shop. But Grice also claims that we should not say of the Arab that he meant that "by the words which he uttered". This suggests that we need to draw a finer-grained picture of the relation between

- (Ia) A meant something by uttering x (on a particular occasion)

and

- (II) A meant something by x (on a particular utterance occasion)

than the coarse equivalence we contented ourselves with back on p. 15. Schematically it seems that in the "Arab merchant" case we are reluctant to pin what he means by his performance onto the token uttered because the merchant does not intend the tourist to recognize his intentions from any features of the token uttered (any other Arabic sentence would have served as well). On the other hand, in the 'cake' example, the speaker does intend the girl to recognize his intention that she should take some cake by noticing certain features of the token uttered (namely, those features that make it a token of the type that she thinks means 'Help yourself to some cake'). So, in this second case, it is natural to pin the speaker's meaning to the token he utters, and say that on this occasion of utterance, he means by the sentence 'Help yourself etc.'. To clarify the relation between (Ia) and (II), therefore, we need to consider the role that the recognition of features of the token uttered plays in a given communication situation. Grice makes a few relevant comments in his 1969 paper, to which I refer anyone interested. But this is a point of fine detail, not of fundamentals.

To sum up on Searle so far: On the natural reading, his 'American soldier' example isn't a counter-example to Grice, for it doesn't satisfy the third condition of (G). If we build up the story so that it does fit the right-hand side of the equivalence (G), we still don't get a counter-example - for now the story fits the left-hand side of (G) as well! The most we can squeeze out of Searle's example is a ground for denying the crude equivalence of (Ia) and (II), and for replacing it by a more complex relation: this is illustrated less confusingly by Grice's parallel 'Arab merchant' example.

So, Searle's example does not amount to much; yet on the basis of it, at least in part, Searle proposes an analysis of 'A meant something by x' which departs from Grice's analysis (G) in two ways (cf. Speech Acts pp. 49-50, or "What is a Speech Act", p. 46 in the reprint). Firstly, Searle maintains that the effect that one intends to produce in one's audience when one says something and means it is always understanding - i.e. recognition of what speech act is being performed by the speaker.

And secondly, Searle holds that the speaker who utters X and means it intends that his intention to produce this understanding in his audience be recognized in virtue of the audience's knowledge of the semantic rules governing X

Now, Searle's second amendment involves a drastic departure from Grice's account, which would not be warranted by the 'American soldier' argument even if the latter were sound. For the whole point of Grice's original analysis was to give an account of 'A meant something by x' which did not involve prior understanding of the notions of conventional meaning, semantic rules or the like. As Grice says, "I would like, if I can, to treat meaning something by the utterance of a sentence as being only a special case of meaning something by an utterance (in my extended sense of utterance), and to treat a conventional correlation between a sentence and a specific response as providing only one of the ways in which an utterance may be correlated with a response". On Searle's analysis, however, the notion of conventional meaning is presupposed by his analysis of 'A meant something by x' - indeed, on his analysis one apparently cannot mean something except by a token governed by semantic rules, which is surely false. If we accept Searle's emendation of (G), then obviously Grice's programme is doomed to circularity.

Well, what reason does Searle offer us in support of his revisions of (G), apart from the discredited "American soldier" example? In the case of his second, more important revision, the answer seems to be: None (or at least, no good reasons). Searle speaks of the need to "reformulate the Gricean account of meaning in such a way as to make it clear that one's meaning something when one says something is more than just contingently related to what the sentence means in the language one is speaking". This suggests that we can't "make it clear that ..." without reformulating the Gricean account in some such way as Searle's. And this suggestion is false - as we saw above when dealing with Quinton's objection. To repeat: A can only have the intention needed to satisfy (G) if he avails himself of means he can reasonably rely on to effect his intentions, and this will of course standardly involve availing himself of the resources of a natural language which he shares with his audience.

What about the first of Searle's revisions of (G) according to which the effect that one intends to produce in one's audience in communication situations is standardly understanding? Well, let's just pause to get our bearings: at the moment we are basically interested in objections to the sufficiency of Grice's analysis, objections of the kind "Conditions (i) to (iii) are satisfied in this case; but this isn't a case where A means something". To say that we standardly want not to produce belief or intention to act but merely understanding is an objection to the necessity of Grice's analysis, an objection of the kind 'there are lots of central cases of meaning, where conditions (i) to (iii) are not satisfied'. Still, since we are already discussing Searle, let's digress, and discuss his claim that Grice's conditions are not necessary conditions of meaning,

that all that is required for meaning is the intention to produce understanding. Of course, his claim rests on an obvious truism: patently, one intends what one says to be understood. However, according to (G), if one says something and means it, then one not only intends to be understood but also (i) intends to produce in one's audience some response such as a belief, and (ii) and (iii), intends this response to be produced by the audience's recognition of intention (i). And it is the latter which Searle is concerned to deny; he puts forward an argument for the case when the intended response is a belief, designed to show that one does not in general suppose that the knowledge in one's audience that one intends to produce a certain belief in them will be one of their reasons for adopting this belief. He cites the case of books on philosophy, and says that an author would be absurdly egocentric if he supposed that his readers would believe what he said merely because they recognized that he intended them to. But as it stands, this is merely an ignoratio elenchi. Grice hasn't claimed that recognition of one's communication intention need be the sole reason that the audience is intended to have for manifesting the required response. The claim is only that such recognition be at least part of the audience's reason for producing the intended response. So it would be quite consistent for Grice to agree with Searle that the author of a philosophy book does not suppose that his readers will believe what he says merely because they recognize that he intends them to. There in fact turns out to be an important point lurking hereabouts, which we will have to re-examine later: but Searle's discussion badly fumbles it.

It is worth emphasizing (to quote Jane Heal in an unpublished paper) that in any case,

*Books on philosophy are quite exceptional in that they purport to present a priori truths as the conclusions of arguments and the author would not have achieved his purpose if his readers did not accept what he said 'for its own sake'. The case of memoirs or history is very different; here it is difficult to see what reason one could have for believing what was said (in the absence of other and independent grounds except one's recognition that the author intended that one should believe it - i.e. was not presenting the narrative as fiction".*

Searle would probably answer that one believed what the author said because one recognized that the author had stated it, and "given an undertaking to the effect that it represented an actual state of affairs" (cf. Speech Acts p.66). But now Searle owes us an account of "giving an undertaking ...": what is "giving an undertaking ..." except putting oneself in a position where one can be reproached if what one says turns out to be false? But one can't put oneself into such a position merely by saying that one does - Jane Heal again:

*To say 'I give an undertaking that p is true' in circumstances where no-one is interested in the truth of p is not to lay oneself open to blame if it turns out false, anymore than promising to do what one's audience does not want creates for one an obligation to do what one has said. It seems therefore that we should regard the blame we often receive if we say things that are untrue as in fact supervenient upon another action we have done, namely making (or trying to make) someone believe something that is untrue.*

If this is right, then the notion of 'stating' that Searle appeals to is itself to be elucidated through the concept of a Gricean communication intention to produce a belief in the hearer.

In this section, then, I have argued that Searle's radical objections to Grice do not work; though as I hinted before we will see that a Searlean objection can be reconstructed which reveals an important defect in Grice's work. But as it stands, Searle's discussion is a rag-bag of more or less interesting failures. We will turn next to examine some objections to Grice's analysis (C) which really do work. They deserve a unit to themselves.

October 1980.