The Meaning-Nominalist Strategy

Jonathan Bennett

[From *Foundations of Language* 10 (1973), pp. 141–68.]

1. Introduction

What did King Charles mean by 'Remember!' on that famous occasion? What does 'Remember!' mean in English? One question involves occasion-meaning, the other conventional meaning; and there are problems about how the two concepts are related. 'Meaning-nominalism' could stand for a pair of theses about this. (1) The concept of conventional meaning can be elucidated through that of occasion-meaning, and this order of explanation cannot be satisfactorily reversed. (2) Cases of meaning need not in any way involve conventional meaning: someone who utters something giving it a certain meaning need not be conforming to any convention for utterances of that kind, nor need he be flouting or extending or launching or trying to conform to a convention.

Any language, properly so-called, must in some way involve conventions; and indeed the concept of language involves that of convention. So thesis (1) entails that a certain meaning-concept can be explained without using the concept of language, and (2) entails that not all meaning is linguistic. For this reason, 'meaning-nominalism' is a better name for the pair than the rival 'linguistic nominalism'.[1]

Meaning-nominalism points to a certain strategy. It suggests that the way to get clear about the concept of language, or about the concept of conventional meaning, is to start with a genus *meaning* and then proceed through the differentia *convention* to the species *conventional meaning*. I am not sure how widely meaning-nominalism is accepted as theory, but it is certainly not always adopted as strategy: not all meaning-theorists start with meaning and then move on to conventional meaning. This may be because we have seemed to lack any viable account of the genus meaning, but I shall contend that a paper by Grice has changed all that.[2]

Then there are problems about the differentia. I have said that this is the concept of convention, and that the species we are interested in is conventional meaning; but now I want to retreat from this for a moment. It is best to see meaning-nominalism as the view that the meaning on an occasion of an utterance-token has a certain primacy over any sort of meaning-in-general of an utterance-type—the meaning that tokens of that type have

usually or
generally or

[1] See David D. Welker, 'Linguistic Nominalism', *Mind* 79 (1970), 569–580.

[2] H. P. Grice, 'Meaning', *Philosophical Review* 66 (1957), 377–388; reprinted in P. F. Strawson (ed.) *Philosophical Logic*, London, 1967.

by rule or
by convention

or whatever. It is then a further question as to which of these notions, i.e. which sort of meaning-in-general, is needed for an account of what language is. It is widely agreed that merely *usual* meaning does not suffice for this purpose.[1] Consider the two forms of statement 'S means that P in language L' and 'When L-speakers utter S they usually mean that P'. Although the two are somehow connected, they are far from equivalent. For one thing: if S means (only) that P in L, then it will ordinarily be wrong to utter S meaning that Q; but from the fact that speakers usually utter S meaning that P, it follows only that departures from this are unusual—not that they are wrong.[2]

Also, it could be sheer coincidence that L-speakers usually mean by that P, but coincidence could not bring it about that S means that P in L. So we need something stronger than merely usual or general or regular meanings. It must involve regularities; but it must be too strong to be entailed by the existence of sheerly coincidental regularities, and strong enough to entail that given linguistic performances are not just irregular but wrong.

These requirements might be met by the idea of regularities which arise from *rules* or from *conventions*, and each of these concepts has found its supporters in the philosophy of language. The concept of rule, however, generates more heat than light. As for convention: this has come to seem even more obscure, unless so construed that language is not after all governed by conventions;[3] and, while some meaning-theorists have still used 'convention', they have not tried to invest it with any very precise meaning.[4] Recently the whole picture changed. David Lewis has given a games-theoretic analysis of the concept of convention, solving problems whose non-solution has years kept the concept out of favor.[5] Crucially, Lewis shows how a convention can be more than a mere behavioral regularity without having to result from any group's agreeing to anything; and he makes no use of metaphor or of 'tacit agreement' or 'behaving as if they had agreed' or the like. That alone entitles the concept of convention to a reprieve; but Lewis's analysis has other good features also, which give the concept of convention a hitherto unsuspected power to clarify, unify, deepen and explain.

Thanks to Grice and Lewis, then, the practising meaning-nominalist is much better off than he was a few years ago. I shall give evidence for this by deploying the Gricean analysis in an explicitly meaning-nominalist way in Sections 2 and 3 and then in Sections 4 and 5 adding in a variant of Lewis's differentia.

Section 6 will offer brief remarks on the origin of conventions, leading on to a discussion in Sections 7 and 8 of the question of how one can learn about conventional meanings. Although this is an old and familiar question, I think that my stiltedly meaning-nominalist approach to it, helped by the work of Lewis and Grice, will let me ask and answer it more clearly and decisively than has previously been done.

Some writers value Grice's work and use certain aspects

[1] Widely, but not universally—see Paul Ziff, *Semantic Analysis*, Ithaca, 1960, 34.

[2] See Gary Iseminger, 'Uses, Regularities, and Rules', *Proceedings of the Aristotelian Society*, n.s. 67 (1966–7), 73–86.

[3] W. V. Quine, 'Truth by Convention', in *The Ways of Paradox*, New York, 1966, esp. pp. 98–99. W. P. Alston, *Philosophy of Language* Englewood Cliffs NJ, 1964, pp. 56–58.

[4] P. F. Strawson, 'Intention and Convention in Speech Acts', *Philosophical Review* 73 (1964), 439–460. J. R. Searle, 'Introduction' in J. R. Searle (ed.), *The Philosophy of Language*, London, 1971, pp. 1–12. That volume contains a reprint of Strawson's paper.

[5] David K. Lewis, *Convention*, Cambridge, Mass., 1969.

of it, yet jettison its meaning-nominalist emphasis—not by denying meaning-nominalism but by starting with the species instead of the genus. This is true of Shwayder, as Strawson noted in his review.[1] Searle, also, offers a Grice-based analysis of something like the concept of conventional meaning, without preceding it by an account of the generic notion of meaning. Rather than merely omitting the step which would make his procedure a meaning-nominalist one, however, Searle defends his approach by positively criticising Grice's. I shall discuss these criticisms in Sections 9 and 10. Although they are not really directed against meaning-nominalism itself, whether considered as a theory or strategy, they raise at least one question which is important for meaning-nominalism. I shall answer it in Section 11.

2. A Gricean program

In these two sections I shall outline an attempt to develop an extensive analysis of the concept of meaning based on Grice's work. Nothing will be dealt with fully, and some aspects—e.g. the notion of word-meaning—won't even be touched.[2] I aim only to indicate a kind of approach to the concept of meaning, and to create a presumption that it is worth a try.

From Grice I take just this conditional:

> If U utters something intending thereby to get A to believe that P, and relying for the achievement of this upon the Gricean mechanism, then U means that P by his utterance.

Grice speaks of U's 'intention that' the mechanism shall operate, but I think that 'reliance on' it is better.

To say that the Gricean mechanism operates is to say, in part, that A recognizes U's intention to get A to believe that P, and is led by that recognition—through trust in U—to believe that P. Because of certain unwanted cases involving a peculiar sort of *contrived cross-purposes*, we need to add further condition, namely that A has no beliefs of the form: 'U is relying on mechanism M but intends or expects me to think that he is relying on mechanism M′ instead.' For details of the cases which make this extra condition necessary, the reader is referred to their inventors, Strawson and Schiffer.[3] The above condition for barring such cases is taken from Grice.[4]

I affirm only the conditional from intending to meaning, which state sufficient conditions for meaning. Cases which would falsify the converse conditional must eventually be dealt with, of course, but not right away It is better to use the intending-to-meaning conditional first to develop an account of one kind of meaning, and then of one kind of language, and then to regain generality by introducing other kinds of meaning and language as variants on what was initially developed. The alternative is to seek full generality at each stage by insisting upon biconditionals. This is heroic,

[1] D. S. Shwayder, *The Stratification of Behaviour*, London 1965, pp. 287–297. P. F. Strawson, review of the foregoing, *Philosophical Quarterly* (1966), 389–391.

[2] I entirely agree with this: 'Sentences mean what they mean because of what the words in them mean; and, conversely, to know what a word means is to know what difference is made to the meaning of sentences in which it occurs by the fact that that word is used and not others.' R. M. Hare, *Practical Inferences*, London 1971, p. 95.

[3] Strawson, 'Intention and Convention in Speech Acts', at pp. 446–449, or pp. 28–31 in Searle's reprint. I base my allusions to Schiffer on unpublished material by him and published references to him; but I believe they will all be supported by Stephen Schiffer, *Meaning*, Oxford 1972, which I have not seen at the time of writing.

[4] H. P. Grice, 'Utterer's Meaning and Intentions', *Philosophical Review* 78 (1969), 147–177; see pp. 155–159.

and it accords with the traditional idea that an analysis states necessary and sufficient conditions; but despite those two claims on our respect, the biconditional approach to this problem seems on present evidence not to be humanly manageable. Using the intending-to-meaning conditional as spearhead, we can thrust through to the centre of the conceptual area want to occupy; and, once we have arrived there, we can build a base from which to range out and capture the rest of the territory. (I do presuppose that Gricean conditions are sufficient for a central, basic kind of meaning and language.) One might prefer to advance on a very broad front, invulnerable to flanking attacks and not committed to mopping-up operations after the main campaign is over; but, as students of Field Marshal Montgomery's career know, a 'tidy battlefield' is usually associated with a painfully slow advance.

Let us take a tribe of anthropoid organisms, imagine a steady growth in our knowledge of them, and see whether we can—aided by the Grice conditional—become entitled to say things of the form 'By doing X that tribesman meant that P'. To do this, we must first be able to credit them with complex *intentions* of the sort mentioned in the Gricean antecedent; and since those are *intentions to produce beliefs* we must also be able to credit them with *beliefs*. Our basis for all this will have to consist mainly in their (non-linguistic) behavior.

I am working on a detailed theory about how non-linguistic behavior can support working concepts of belief and intention, but there is no space for it here. In the present paper I merely assume that the job can be done—i.e. that non-linguistic behavior can have enough suitable complexity to support all the needed distinctions. Possible objection: 'It isn't just a matter of what distinctions can be drawn. The fact is that we could never be entitled to credit languageless creatures with more than acting *as if* they believed that P or intended to produce T. Non-linguistic behavior cannot support the concepts of belief and intention, but only feeble cousins of them.' If you wish, consider my uses of 'belief' and 'intention' etc. to be suitably enfeebled throughout this paper: the objection will then be met, and everything I have to say will go through as before. That the objection can be met by a mechanical amendment is a measure of its triviality.

However, we have to attribute beliefs about intentions, and also intentions to produce beliefs. This may seem a tall order, even to those who agree that non-linguistic behavior can justify the attribution of intentions and beliefs of simpler kinds. I shall argue that if the short order can be met then so can the tall one. In the following argument, I use 'non-linguistic intention [belief]' to mean 'intention [belief] attributed solely on the basis of non-linguistic behavior'.

The argument has two premises which will, I think, be accepted by anyone who thinks that there can be non-linguistic intentions and beliefs at all. **(1)** Non-linguistic beliefs and intentions are behavioral dispositions of certain sorts; or at least the existence of behavioral dispositions can be sufficient for the existence of beliefs and intentions. **(2)** There can be a non-linguistic intention to produce a certain disposition in something (e.g. to make it flexible or friendly), and there can be a non-linguistic belief that something has a certain disposition (e.g. that it is fragile or afraid).

If those premises are right, then non-linguistic intentions to produce beliefs, and non-linguistic beliefs about intentions, don't involve any difficulty of principle; for they are just intentions to produce (or beliefs about) dispositions which are merely more complex than flexibility and fear.

Throughout this conceptual area, the high-level must yield to the lower-level. For example, behavior which can be well accounted for through something of the form 'Whenever the animal receives an S it performs an R', where S is a

sensory kind of stimulus and R is a motor kind of response, ought not to be diagnosed as behavior performed with an intention on the basis of a belief. A similar rule governs *what* beliefs and intentions we may properly ascribe: the ascribed belief or intention must be as simple, as low-level, as contentless as the data permit; so that, for instance, we ought not to say 'It intended to make me afraid' if the behavior is as well accounted for by 'It intended to make me run away'.

So any attribution of a non-linguistic intention to produce a belief, or of a belief about an intention, is prima facie open to a 'challenge from below'. Whether the challenge succeeds in a given case, though, depends upon the empirical facts of the case. Did the creature intend to make me afraid, or merely intend to make me run away? That depends upon such matters as how it would have acted if I had been visibly unable to run away. This may be hard to discover; but we know how in principle to get such question answered, and so we know in principle how to adjudicate between a high-level psychological diagnosis and a lower-level or otherwise simpler challenge.

It may sometimes seem unlikely that a given high-level diagnosis D could have an adequate behavioral basis. As a start to exploring whether it could, develop a behavioral story which seems to point towards D but which in fact be handled by a lower-level diagnosis: what the creature thinks is not that I am hostile but just that I am going to hit it; what it intends is not that I shall think it is wounded but merely that I shall leave it alone. Then try to vary the behavioral story so as to leave D standing while eliminating the lower-level challenger. Perhaps then some other challenger will be on the cards. If so, develop further variations which will eliminate it too, and so on until you have a behavioral story for which D is the best, simple lowest-level account that can be given. If D cannot be freed from lower-level challengers, then it is something we cannot be entitled to say on the basis of non-linguistic behavior. I do not deny that there are attributions which can properly be made only to creatures with a language. I claim only that attributions of intentions to produce beliefs, and of beliefs about intentions, are not all of this sort. Nor have I undertaken to prove this. I have sought only to render plausible, in a general way, the hypothesis that there could be behavior-patterns which would justify these attributions.

3. A primeval case

Suppose that we have learned a great deal about our subject-tribe's needs and wants and perceptions and beliefs and intentions, including beliefs about intentions and so forth, though we have not yet been in a position to credit them with a language or even with ever meaning anything by what they do. Given this much I want to describe a primeval Gricean case—a situation where we can properly say, for the first time in our knowledge of the tribe, that a tribesman has done something intending to get an audience to believe something, and relying for this on the Gricean mechanism.

Suppose we observe the following: Utterer performs a crude enactment of a man being hit on the head by a heavy object, and Audience sees this and steps out from under the palm-tree just before a coconut falls on the place where he was standing. I am not asking the reader to agree, on an impressionistic basis, that the case has a vaguely Gricean look about it. I want the situation, and our knowledge of the tribe, to be such that a Gricean diagnosis is the best one—the one we ought to adopt just as students of the tribe's behavior, apart from any interest in meaning or in Gricean cases as such. This is a stiff demand, but I think it can be met.

I submit that our knowledge of the perceptions, beliefs, wants and intentions of the tribesmen, and of their knowledge of these aspects of one another, could entitle us to credit Audience with believing that Utterer intended by his performance

(a) to achieve something;
(b) to change his environment;
(c) to change Audience;
(d) to change Audience either visually or in some way arising from the visual change;
(e) to get Audience to believe something;
(f) to get Audience to believe something about someone's being hit on the head;
(g) to get Audience to believe that he risks being hit on the head by a coconut.

What matters is just that Audience thinks that Utterer intends **(g)**. He doesn't reason his way along the sequence, though *we* might do so in considering what Utterer intended or what Audience believes that Utterer intended. My point in gradually building up to **(g)** is to indicate how behavioral data might entitle us to credit Audience with thinking that Utterer has the complex intention **(g)**. Without that break-down into seven steps, it might seem impossible that our evidence should point to **(g)** rather than to some other account of what Audience believes Utterer to intend; but the break-down changes the picture, I think. For example, it isn't hard to see what might justify the shift from **(b)** to **(c)**: Audience can see that Utterer can see that his performance won't get trees felled or fires lit or fish caught, and that in general it hasn't a hope of affecting—in a manner answering to Utterer's interests—anything except Audience himself. Similar considerations underlie **(a)**, and each move from **(a)** through to **(d)**. Nor is there any special problem about the move from **(d)** to **(e)**, except for those who object in principle to non-linguistic beliefs about intentions to produce beliefs. The shift from **(e)** to **(f)**, obviously, depends on the fact that Utterer has acted in a way which could serve as a natural reminder or pointer for Audience to the idea of someone's being hit on the head. The final move from **(f)** to **(g)** rests on Audience's knowing that the only belief about someone's being hit on the head which Utterer could want Audience to acquire at that time would be the belief that Audience himself was in danger of being hit on the head. In the circumstances there just isn't anything else about head-hitting that Utterer could think it worthwhile to try to get Audience to believe. But this works only because we have already got as far as **(f)**—that is because the peculiar nature of Utterer's performance has narrowed down the choice to a certain relatively small genus of propositions. So the move from **(e)** to **(f)** is absolutely crucial. I shall return to it later.

Now, if we can credit Audience with thinking that Utterer intends him to believe that P, then we can also credit Audience with being led by that to believe that P. All we need is a general background of trust which can created without invoking meanings: for Utterer can show that he is generally disposed to give Audience true and helpful beliefs, by taking pains to ensure that Audience has evidence that he might otherwise overlook.

What about *contrived cross-purposes* of the sort mentioned early in Section 2 above? Can we credit Audience with having any belief of the form 'Utterer expects to achieve his goal in *this* way, but he expects me to think that he expects to achieve it in *that* way'? To develop the story along those lines we should need a good head for heights, and a steady hand for the management of subtle and complex behavioral details; but no doubt it could managed somehow. Avoiding such a diagnosis, however, is not a problem. To say that we can develop the story along other lines—i.e. that we

can avoid crediting Audience with thinking that Utterer is contriving cross purposes—is to put it mildly.

So we can be entitled to say that what operated in Audience was the 'Gricean mechanism' as defined in Section 2. Audience thought that Utterer intended to get him to think that P: he didn't think that Utterer expected this to be brought about through Audience's being in error about an aspect of what Utterer was up to; and his recognition of Utterer's intention led him—through trust in Utterer—actually to believe that P.

But our entire basis for all these attributions to Audience consists in behavioral data which have been available to Utterer also. So there is no difficulty in principle about our being able to say not just that the Gricean thing happened but that Utterer intended it to happen. That is *Utterer intended Audience to think that P, and he was relying on the Gricean mechanism for the realization of this intention.* So we have our first Gricean case—our first case, if Grice is even half right, of a tribesman's doing something and thereby meaning that P. We have had to appeal to general facts about the tribesmen's behavior-patterns, standing wants, dispositions to believe, and so on, but not to any other cases of meaning. *A fortiori*, we have not had to bring the given case under any general meaning-convention. All this, I submit, is a substantial vindication of meaning-nominalism as not just a pious hope but a practical analytic program.

Let me emphasize: this section claims that *we could have* behavioral data which entitled us to say that a tribesman's action was a case of meaning according to Grice's analysis. I have tried to make this claim plausible by indicating the sorts of data which would support various parts of a total Gricean diagnosis of an action; but I have not tried to present in detail a possible set of data which would do the job.

4. Convention

Lewis's analysis, in essence, is as follows.[1] A group may need to coordinate their actions in some way, so that what it would be best for each to do depends on what the others do. If they wish to meet, each wants to go were the others go. If they wish to re-establish the broken telephone connection, each wants to dial if and only if the other hangs up and waits. Coordination may be reached through prior agreement, either about the particular case or about all cases of a certain kind; but there are also other bases on which each member of the group may guess what the other(s) will do. In particular, it may be that a certain coordination-achieving procedure C has regularly been followed by the group in situations of kind K, and that there is mutual knowledge within the group that this is so—meaning that each member knows that in K situations C has been followed, knows that the others know, knows that they know that he knows, and so on. In a new K situation, if the group cannot convene and discuss how they are to act, they are almost certain to follow C once more. This is because, unless some extraordinary event has created a mutual belief among them that some other procedure is uppermost in everybody's mind, C is the only procedure which each can reasonably think the others are pinning their hopes on, and thus the only one that each can pin his own hopes on. This being so, it will be reasonable for each to perform the action which is his part of C, and so it will be reasonable to expect that C as a whole will be followed by the group.

This is a situation where a behavioral regularity—*in K situations C is followed*—is maintained because it is mutually known to have obtained in the past, and provides the group

[1] Lewis, op.cit., p. 58.

with their best or only chance of achieving coordination in new K situations. A regularity of that sort, maintained for that reason, is a convention.

I borrow 'mutual knowledge' and 'mutual belief' and their cognates from Schiffer (op.cit.). Mutual knowledge, in his technical sense, can easily be derived by extension from the two-person case, which is all that needs to be discussed here. To say that *x and y mutually know that P* is to say that in the series of propositions

x knows that P
y knows that x knows that P
x knows that y knows that x knows that P

and so on, and in the series

y knows that P
x knows that y knows that P
y knows that x knows that y knows that P

and so on, there are no false members. Perhaps every member of each series can be true; but arguably each series degenerates into a sort of meaninglessness, its later members being neither true nor false but mere word-play. Cargile has argued persuasively that after about the fourth member of the series there cannot by any serious question of truth or falsity.[1] I need not go into the point. The negative formulation, which I have adapt from Grice's handling of the contrived cross-purposes mentioned in Section 2, says all that I need and keeps me out of controversy.[2]

In trying to apply Lewis's account of convention to the concept of meaning we immediately encounter a problem, namely: in cases of meaning, what is being coordinated with what? Lewis's whole account is tied to the coordination of actions; but in cases of communication it is not clear what actions speaker and (especially) hearer are in question. 'Each "gives a meaning" to the utterance, and coordination is achieved if they give the same meaning'—perhaps, but that formulation just papers over the cracks, for meaning-giving is not a kind of action.

Lewis applies his concept of convention smoothly to one sort of communication—called 'signaling'—which does involve the coordinating of actions, properly so-called, by speaker and hearer. But he points out that signaling is very restricted and specialized: it requires that there be some action which both speaker and hearer judge to be the best thing for the hearer to do if the communicated message is true; and there are other severe limitations as well.[3]

This leads Lewis to emphasize a quite different application of the concept of convention, in which conventions serve to coordinate speaker with speaker. The vital question here concerns the relation between U_1's speaking and U_2's speaking, not between U's speaking and A's understanding. Lewis says:

> Generality is served by concentrating on the one-sided coordination among communicators which is present in all indicative communication, not on the occasional two-sided coordination between a communicator and his audience.[4]

[1] James T. Cargile, 'A Note on Iterated Knowings', *Analysis* 30 (1969–70), 151–155; also Colin Radford, 'Knowing and Telling', *Philosophical Review* 78 (1969), 326–336.
[2] Grice, 'Utterer's Meaning and Intentions', p. 159.
[3] Lewis, op. cit., pp. l60ff.
[4] Ibid., p180.

No doubt coordination among speakers is vital to all real-life linguistic situations: real languages are social, in the sense of being used by many speakers, and the question of whether you and I can communicate with one another is often practically equivalent to the question of whether we speak the same language. But generality of an important kind is lost if we attend only to social languages. There are no conceptual difficulties, though there may be psychological ones, about the idea of a community in which no-one's idiolect is significantly similar to anyone else's, though each idiolect is understood by the whole community.

That possibility, however improbable it may be in our world, highlights the fact that communication is essentially a transaction between speaker and hearer; and it seems plausible to suppose that conventions are somehow relevant to many cases of communication, considered just in themselves as speaker-hearer transactions and independently of any social language in which they may be embedded. So I want if I can to adapt Lewis's analysis of convention so that it will apply to two-person communication without the unduly severe restrictions imposed by 'signaling' in Lewis's sense.

The problem about doing this is perhaps obvious. In approaching communication from the direction of Grice's analysis of meaning, I make the hearer's essential role one of acquiring certain beliefs; and, as Lewis remarks, the acquiring of a belief 'is not normally a voluntary action and hence not an action in conformity to convention' (Ibid.). I agree with this, but I believe that Lewis's account of convention can be widened so that it applies not only to the species *action* but to a genus of items which include not just actions but also belief-acquisitions and perhaps some other items as well.

This generalization of Lewis's account is a fairly long business, and I shall devote my next section to it. I think it is instructive enough to be worth the trouble, and Lewis tells me that he thinks so too. For the purposes of the present paper, however, we could save ourselves that trouble by adopting a formulation borrowed from Schiffer (op.cit.). Let us say that there is a convention whereby utterance-type S means that P within a given tribesman's idiolect if (a) in the past he has uttered S only when he meant that P, and (b) this fact is mutually known to him and his hearers, and (c) because of the mutual knowledge mentioned in (b) it continues to be the case that when he utters S he means and is understood to mean that P. This formulation, like Lewis's, embodies the crucial idea of conventional meaning's being more than mere usual meaning, the extra element consisting in the idea of a regularity's being adhered to because it is mutually known to have obtained in the past—where 'because' reflects a reason rather than a cause. Much of the detailed work Lewis does on the basis of his analysis—including his relating 'convention' to 'arbitrary', to 'agreement', to 'norm' and to 'rule'—could be fairly easily reconstructed in terms of Schiffer's formulation.

5. Lewis's analysis generalized

Satisfactory as that short-cut is, I think that there can be profit and pleasure in seeing how Lewis's analysis can be generalized so that it directly covers coordination between speaker and hearer. The essence of this generalizing procedure is a move from the species *action* to a wider genus of what I shall call *doings*. When I use parts of the verb *to do* to stand for this genus, shall use italics as a reminder that this is a term of art.

When I say that a person *does* something at a certain time, I mean that

> there is some value of ϕ such that (1) that person ϕs at t, and (2) one's beliefs can be at least partial reasons for one's ϕing.

On this account human actions are all *doings*. So also are belief-acquisitions: (1) they can be reported in sentences of the form 'x ϕs at t', and (2) the reasons for one's coming to believe that P can include other beliefs that one has. Also, one can deliberate, consider, wonder what to believe (and whether to believe P), just as one can deliberate etc. what action to perform (and whether to perform X). In each case one can weigh up reasons, and these will be—or at least include—beliefs. *Doings* include the acquisition of various states which have a belief-component, e.g. fears, and perhaps other items as well, but none of these have to be considered here. Nor need I follow up the fact that my explanations allow us to classify as *doings* items which are not happenings at all, e.g. beliefs. The doings which centrally concern me are all actions or belief-acquisitions, which are happenings. '

I contend that the class of *doings*, as I have explained it, is a conceptual natural kind rather than just the product of a pun. I defend this not by minimizing differences but by stressing similarities. Reasons for actions, unlike reasons for beliefs, typically include not just beliefs but also wants and approvals and aversions and the like. (The wish may be father to the thought, but not a reason for it.) I don't deny that there is this difference. Still less do I represent belief-acquisition as a kind of action. Someone might voluntarily act to acquire the belief that P because having it would help him to behave in money-making or God-pleasing ways, say. His reasons for doing this would include wants or the like, and any beliefs which they included would not tend to support P but merely help to make the practical case for acquiring the belief that P. I identify this very special case only in order to set it aside. I am concerned with the standard case in which 'his reasons for acquiring the belief that P' are 'his reasons for thinking that P' or 'the beliefs of his which support P or tend to confirm that P is true'. In this standard case, wondering whether to believe P is wondering whether P.

Reasons for acting and reasons for believing, despite their differences, have enough in common for my purposes. The vital point is that what someone will come to believe, just like what actions he will perform, can depend upon and be predicted through what beliefs he now has. The importance of this will appear in a moment.

Lewis's account of coordination and convention can be rewritten with 'action' replaced by the more general '*doing*', and *mutatis mutandis*. It will then cover not just the cases to which Lewis's original analysis applies but also others including this:

> In a community of two people, U wants to act so as to get A to believe that P, relying on a mechanism which involves A's recognizing that intention; and A, confronted with U's action, wants to know with what intention it was performed. In this situation, 'coincidence of interest predominates' (Lewis, op.cit., p. 14), and coordination is achieved if U does act with the stated intention and A does recognize what that intention was, i.e. comes to have a true belief about why U acted as he did. Now, if U and A mutually believe that in the past whenever U has performed an X action he has meant and been understood to mean that P (all that being spelled out à la Grice), then this mutual belief *both* gives U a strong reason for again performing an X action as his way of realizing that intention *and* gives A a strong reason for believing that when U performs his next X action it is with the intention of getting A to believe that P. And so a certain coordination in their doings is achieved: their mutual belief in or knowledge of a past regularity has led them to maintain the regularity in further instances,

> regarding it as their best chance of continuing to achieve coordination.

The vital point is that the relevant *doings* in this situation, although not all actions and not all "voluntary', are all under the control of beliefs in such a way that they can be explained as arising from the existence of a certain mutual belief and can be predicted accordingly.

Cases of communication differ from Lewis's examples in four ways which do not consist in—though some may result from—the fact that communication involves *doings* which are not actions. Striking as these differences are, they do not vitiate my procedure.

(i) The utterer's question is: what can I do which will communicate that P? He has a certain intention, and his problem concerns what action to slide in under it, so to speak. This can be seen not as a practical problem of how to act but as a purely theoretical problem about what action is most likely to realize the given intention. One merit of my concept of a *doing* is that it spares us from having to delve too deeply into this matter. The utterer's problem, whether of the form 'How to act?' or 'What to believe?' or a peculiar mixture of the two, certainly has the form 'What to *do*?' in my technical sense of '*do*'; and that is all I need.

(ii) The hearer's *doing* depends logically upon the speaker's, for what the hearer *does* is to acquire a belief about the speaker's utterance. This seems to be connected with the next point.

(iii) The speaker has a coordination problem because he wants to realize a certain intention, and it is only after he has tackled the problem—i.e. made his utterance—that the hearer has a coordination problem at all.

Perhaps features **(ii)** and **(iii)** are present only because we have generalized from actions to *doings*—I am not sure about that. Anyway, the generalized Lewis analysis can take them in its stride. For it remains the case that speaker and hearer want to produce a state of affairs in which their *doings* are related in a certain way, and that they can be helped to achieve this by their mutual belief that a certain regularity has obtained in the past.

(iv) If speaker and hearer mutually believe that whenever the speaker in the past performed an X action he has (for short) meant that P, this will suffice to get them to do what is needed for coordination in this new situation That is, they need only a mutual belief about a past regularity involving only one of them. This, however, is not because of the move from 'action' to '*doing*', but merely because we are dealing with coordination amongst just two people. To take an 'all action' case with only two people: if you and I mutually believe that in the past whenever our phone connection has broken I have hung up and waited, that is enough to give us grounds, the next time it is broken, for me to hang up and wait and for you to redial.

In communication the hearer doesn't 'act according to a convention', because he doesn't act at all. But such cases are governed by conventions: the hearer's understanding is explained by his knowledge of a convention, i.e. by his being party to mutual knowledge of a past regularity satisfying certain conditions.

Before leaving Lewis, I should mention something I have so far suppressed. It is that a statement of the form 'In doing X, U adheres to regularity R as a convention' may entail that U means something by X, even if R itself involves no reference to reliance on the Gricean mechanism or indeed to anybody's intentions. Lewis's proof of this is based, of course, on his analysis of convention (Ibid., pp. 154–156); but my generalization of the analysis doesn't affect the essentials of the proof, and merely serves to bring more regularities within its scope. The essential point is that the statement 'When

X did U he was adhering to R as a convention' is a complex statement about the beliefs and intentions of U's which explain his doing X. Furthermore, it involves attributing to U just the sort of thing that is so characteristically Gricean, namely second-guessing about someone else's beliefs and intentions regarding U. For details, see Lewis.

Lewis agrees that Gricean meaning can occur without conventions' being in any way involved (pp. 158–9); and I have chosen to start there, introducing convention at a later stage. Still, it does need to be noted that if 'meaning' obeys the Gricean conditional, and 'convention' obeys the Lewis analysis, the phrase 'conventional meaning' is pleonastic.

6. Origin of meaning-conventions

If a convention is, in brief, a regularity which the tribe maintain because they know it to have obtained in the past, then its first instances cannot have occurred as instances of a convention. Why, then, did they occur? As Lewis shows in an illuminating discussion, there could be many answers to this (pp. 83–88). A convention could originate in an agreement: the parties agreed to a certain regularity, conformed to it (for a while) because they had agreed, and (thereafter) because it was established and mutually known and the best available solution for a certain kind of coordination problem. This is an unlikely account of the origin of meaning-conventions, however, even if Lewis is right that an agreement need not involve the use of language (pp. 87f).

A more plausible guess is that meaning-conventions evolved from regularities whose instances were like my primitive Gricean case, with meaningful utterances becoming more conventional in their basis as they became less able to stand on their own feet. To give a simple illustration: the second time Utterer wants to warn Audience about a coconut, his enactment of a man being hit on the head is even cruder than it was the first time, so that it might have failed to carry Audience's mind to the thought of someone's being hit on the head, if Audience hadn't recalled the earlier coconut-warning—this being exactly what Utterer intended should happen. This is not yet a convention, but it is close.

My meaning-nominalist program does not require me to opt for any one theory about the origins of meaning-conventions, just so long as it allows that such conventions could originate somehow.

7. Knowledge of meaning-conventions

If someone is to mean that P by an utterance without any meaning-convention's being involved, then he must regard his utterance as somehow naturally connected with P—for example by closely resembling something which would be strikingly characteristic of any state of affairs in which P obtained, or by resembling something which reminds the audience of some particular state of affairs where P held, or the like. Thus, in my origin Gricean case, the utterer enacted a man being hit on the head by something like a coconut.

This is obviously tremendously limiting, and one great service convention renders is to remove those limits. In every case of meaning it must be reasonable to expect the audience to connect the utterance with P rather than with any other proposition; and if this has to be achieved by the utterance's being a natural reminder of P, so to speak, that severely limits what the utterance can be like. But with conventional meaning the utterance has only to be of some noticeable kind which regularly means that P; and that imposes no restrictions on what physical features it may have. Any sort of utterance could conventionally mean that P; and once such a convention exists and is known to the tribesmen, individual utterers can reasonably expect instances of it to be understood by their audiences. (That point holds good

even if only one tribesman ever utters anything: it concerns primarily speaker-hearer conventions, and only derivatively speaker-speaker conventions. That is one reason for wanting the concept of convention to help with idiolects and not just with social languages used by many speakers: the removal of the cramps of the 'natural reminder' requirement doesn't logically require that more than one speaker be involved.)

Leaving aside the question of how conventions get started, let us ask: given that a meaning-convention is operative within the tribe, how can anyone—young tribesman, or we as observers—come to know about it? To learn that S conventionally means that P we must learn that usually when an S-token is uttered the utterer means by it that P, and so we must be able to discover in some instances that individual utterances of S-tokens meant that P. But how can we understand any individual instance of the convention without knowing the convention? This is a real problem, because an utterance-type which conventionally means that P is unlikely to have tokens which portray or enact or naturally remind one of (anything associated with) P. In short, an utterance is unlikely to relate to its meaning both conventionally and in some other way; yet the learner has to connect them while not knowing the convention. How is the trick worked? Some conventions can be verbally explained, but obviously that cannot be the whole answer. We must learn about some meanings just by observation of the tribe's ordinary communicating activities.

It doesn't matter that 'for most utterances the nonlinguistic context simply does not make obvious some purpose the speaker has in uttering it',[1] just so long as there are enough utterances for which the speaker's purpose can be known without reliance on knowledge of meaning-conventions. In fact, not even that is needed, for no case needs to be fully diagnosable or understandable in complete isolation. Suppose we want to test the hypothesis that utterance-type S conventionally means that P. Even if we can never say of any individual S-token that given all the circumstances it must have meant that P, we may still learn that one S-token meant something within a certain range, a second meant something within a different but overlapping range, and so on—until eventually there is only P left as common to all the cases that have been studied. (This is subject to fuzziness in the notion of propositional identity; and if Quine's indeterminacy thesis is right there will also be another challenge to P's uniqueness.[2] But these are complications, not objections; and in what follows I shall ignore them.) If there is no convention governing S-tokens as such, then every hypothesis about the meaning of S will fail. I use 'hypothesis' advisedly: the learner need not be unaware that there are meaning-conventions operative within the tribe. Indeed, he may have an easy clue as to when an utterance has a conventional meaning—assuming that the tribe will have some physical kind of behavior, e.g. vocal behavior, which is their main vehicle of meaning and is seldom engaged in for any other purpose.

It is plausible to suppose that one can learn about individual conventions by this sort of elimination. However, I want not just plausibility but a recipe for generating evidence about conventions. That is, I want kinds of facts which are often available and which enable us, as learners, to identify or at least circumscribe a speaker's meaning without knowing what convention he is applying. But there is a danger to be averted, namely that these kinds of facts—if we can find them—will be available only in cases where the convention

[1] Welker, op. cit., p. 571.

[2] W. V. Quine, *Word and Object*, New York 1960, ch. II.

is somewhat idle. The more surely and narrowly we can circumscribe the speaker's meaning without knowing what the relevant convention is, less essential work the convention will be doing in that situation—or so one might think. For if we can understand the utterance without knowing the convention, why shouldn't the primary or intended hearer do likewise? It would be depressing if meaning-conventions had to be learned through attention to the weakest examples of conventional meaning, where meaning-conventions are not entirely needed for communication to succeed. The problem would be solved if there were enough cases where the language-learner knew something relevant which the primary hearer didn't know because he was deaf or looking the wrong way or the like. That line of solution, however, threatens to put language-learning too greatly at the mercy casual circumstance. Do we just have to hope that we shall find enough situations where we know something helpful which the primary hearer doesn't know? My own solution to the problem, to be expounded in the next section, includes something like the solution sketched above, buttressed by a complete answer to this latest difficulty. Specifically, I shall show that it must be the rule rather than the exception for the primary hearer to be ignorant something relevant which could be available to the learner; so that, rather than merely having to hope that enough such cases will occur, we can depend upon their doing so.

Any theorist about meaning ought to face this problem about how one can learn about meaning-conventions, for it can hardly be controversial say that we need access to occasion-meanings other than through knowledge of conventions. It is sometimes said that we have such access through observations of the circumstances in which various utterances are uttered: that is no doubt true, but it doesn't say much except that the problem's solution, if it has one, lies in the public domain. I want something more specific, preferably not a mere inventory of dozens of kinds of clues to meaning, even ones united by a 'family resemblance'.

There is a worthwhile intermediate level of generality. I shall describe two sources of knowledge about meanings: although much more specific than 'observation of circumstances', they are still highly general. Also, they are central and fairly comprehensive: many of our sources of knowledge about meanings are special cases of these two.

8. Hearer's belief and hearer's need

In each speaker-hearer situation, according to the Gricean paradigm, the speaker has a fair degree of expectation that the hearer trusts him. This, together with a weak further assumption to the effect that the tribesmen are both attentive and retentive, implies first that the hearer does generally trust the speaker and second that the speaker is generally to be trusted—i.e. seeks to give the hearer a true belief. That involves his not being deceitful and not being relevantly misinformed (I ignore the case where deceit and misinformation cancel out, yielding truth), and that may seem a lot to swallow. But the assumption about deceit is supported by the general reflection that a public language is a cooperative endeavor which cannot function without a measure of goodwill. Lewis's analysis points to a deeper grounding for this same point, for according to it a convention is essentially tied to situations which are non-competitive in that 'coincidence of interest predominates' in them. The assumption about the speaker's not being relevantly misinformed can also be defended on other grounds than that we need it for the Gricean paradigm. If we are not just frivolously making mysteries, the thought 'Perhaps the speaker is usually mistaken in the beliefs he tries to communicate through the Gricean mechanism' must

be allowed to stand or fall with the stronger thought 'Perhaps the speaker is usually mistaken in his beliefs'; and I think it will be generally agreed that the latter thought can be dismissed—and therefore the former with it. There remains the question of what is wrong with the idea of someone's being mostly wrong in his beliefs.[1] I suspect that the answer lies in the facts about the behavioral basis for attributing non-linguistic beliefs. A full account of that basis would, I suspect, contain as a well-grounded theorem the claim that it must be the exception rather than the rule for the attributed belief to be false.

So I assume that we have a large variety of paradigm Gricean speaker-hearer situations, i.e. ones in which the hearer rightly trusts the speaker. I now turn to the promised two sources of knowledge about what speakers mean.

One of these is an item which the hearer cannot have in time to use it as a substitute for knowledge of the relevant meaning-convention. This is the evidence provided by the belief the hearer does in fact acquire as a result of hearing what the speaker says. Although when the episode is over the hearer too can know what belief he acquired, it is obvious that since he acquired it through understanding the utterance his knowledge of it can't be a means to his understanding the utterance, and so can't threaten to render the convention idle on this occasion.

This hearer's-belief basis for understanding utterances is available to us only when the hearer (1) understands the utterance, (2) believes it, and (3) manifests his belief in action. There is no problem about (1). That hearers often understand speakers is implied by the very question we are asking, namely the question about how one can learn the content of the tribe's operative meaning-conventions. The assumption (2) is a structural part of the Gricean paradigm presented above. As for (3): the whole program—including everything that leads to the asking of our present question—presupposes that a good proportion of the tribesmen's beliefs are manifested in their behavior.

The second source of evidence about meaning is one which can be used by us as learners but can't normally be used by the primary or intended hearer. This consists in evidence as to what the speaker could want the hearer to come to believe at that time in those circumstances. Although we shall often have no evidence about this, there will be many cases where we can in this way circumscribe the speaker's meaning, sometimes narrowing it down to virtually a single candidate (subject only to the inevitable fuzziness and—if Quine is right—to rivals which are ineradicably on the cards and therefore do not harm my account). To take an example at the extreme of simplicity: if the hearer is threatened with a falling coconut which he can't see but the speaker can, and if he is in no other danger and not in a position to help anyone else, then this is evidence that the speaker means by his utterance that the hearer is threatened with a falling coconut.

Of course the hearer might look up and see the coconut. In other cases it will be harder for him to discover independently what it is that the speaker wants him to believe—e.g. because it involves features of the situation that he hasn't the skill or knowledge to detect (or the opportunity: the vital fact might concern something that happened earlier, in the hearer's absence). In any case, if he does discover that the speaker is likely to want him to believe that P, and comes to understand the speaker's utterance on that basis rather than through knowing the relevant convention, then the

[1] See A. M. MacIver, 'Knowledge', *Proceedings of the Aristotelian Society*, suppl. vol. 32, 1–24, especially pp. 23–4.

utterance has failed not just as conventional meaning but as meaning. I shall try to explain.

In the paradigm which I defended early in this section, discovering what the speaker wants the hearer to believe is discovering what the hearer needs to believe or would be helped by believing; and, with only negligible exceptions, discovering *that one needs to believe that P* involves discovering *that P*. If the hearer finds that he needs the belief that P, that will normally be by coming to believe that P. But if that is his basis for understanding the utterance, then he can't be led by that understanding to the belief that P—for ex hypothesi he has that belief already. In such a case therefore the intended Gricean mechanism has been short-circuited, so radically indeed that the speaker's utterance has nothing at all to do with the hearer's coming to believe that P (except perhaps trivially, e.g. by making a noise which causes the hearer to look up and see the coconut). In some cases, a short-circuit could easily happen, while in others it couldn't. But to the extent that a given speaker in a given situation can reasonably mean that P, thus relying on the Gricean mechanism for the production in his hearer of the belief that P, to that extent a short-circuit is off the cards in that particular situation.

In saying that the hearer's-need basis is not normally available to the primary hearer in a meaning-situation, therefore, I am not making a wild statistical guess. Rather, I am saying something that follows from the essence of meaning. If in a given case the hearer's-need basis is available to the hearer, then the speaker need not rely entirely on the Gricean mechanism; so he relies on it either wrongly or not at all; and either way the case is less than a paradigm of meaning.

Objection: 'If the hearer cannot easily discover that P independently of the speaker's utterance, why assume that a learner will be able to do so?' Well, if the hearer cannot easily discover that P independently of the speaker's utterance, why assume that *the speaker* will be able to? The answer to this is just that there are in many cases relevant differences of skill, knowledge and epistemic opportunity between speaker and hearer; and there is no reason why many of these should not also obtain between eavesdropping learner and primary hearer also. I have no unified over-all account of what these differences are, but my argument above shows that if they didn't often obtain between speaker and hearer, it would seldom be appropriate to utter something and mean something by it.

To sum up. The hearer's-need basis for discovering what the utterances mean is, though frequently available to us, not normally available to the hearer, even though the hearer is normally someone just like us. That makes it a good source of knowledge about meaning-conventions: it enables us to discover what a given utterance means without our knowing what convention governs it; yet it does not significantly threaten to imply that the convention could easily be idle in that particular situation; and we have a principled account of why it has both these desirable features.

9. Understanding and convention

Searle has complained of inadequacies in Grice's approach to meaning,[1] and has used these complaints to justify replacing 'the original Grice analysis of non-natural meaning'

[1] J. R. Searle, 'What is a Speech Act?' in Max Black (ed.), *Philosophy in America*, Ithaca 1965, pp. 221–239, reprinted in Searle (ed.), *The Philosophy of Language*, pp. 39–53; *Speech Acts*, Cambridge 1969, pp. 43–50. All my quotations from Searle are, unless otherwise noted, from pp. 43–49 of *Speech Acts*, though some of them also occur on pp. 228–230 of 'What is a Speech Act?' or pp. 44–46 of Searle's reprint thereof. For Searle's revised analysis, see pp. 49–50 of *Speech Acts*.

by 'my revised analysis of the different concept of saying something and meaning it'—an analysandum which Searle also expresses as 'uttering a sentence and meaning it' and as 'the performance of an illocutionary act in the literal utterance of a sentence'. This analysandum is in fact the concept of uttering something and meaning by it what one thinks it conventionally means.

In disentangling Searle's points against Grice, let us start here:

> On Grice's account it would seem that any sentence can be uttered with any meaning whatever, given that the circumstances make possible the appropriate intentions. But that has the consequence that the [conventional] meaning of the sentence becomes just another circumstance.

On the face of it, it seems right to say that in the Gricean type of account which I have been presenting, an utterance's conventional meaning is just a 'circumstance' which affects what a speaker can intend by his utterance. Searle apparently wants an account in which the speaker's intention is not merely affected by conventions but is *about* conventions—i.e. he wants the speaker to be allowed to intend that the conventions governing his utterance be recognized by his hearers.

Now, the Gricean approach can handle some kinds of intentions about conventions, namely ones in which the speaker intends that the hearer acquire a certain belief through recognizing the conventional meaning of the utterance. That involves something of the form 'He intends that a certain mechanism (involving recognition of conventions) be instrumental in the realization of his primary intention'. I contend that there is no significant difference between *intending that* mechanism M be instrumental and *relying on* mechanism M for the realization of one's primary intention; and of course the Gricean approach can easily allow that sometimes the speaker relies on the hearer's knowing the conventional meaning of the utterance.

However, as well as claiming generally that speakers can have intentions about conventions, Searle makes a more specific claim which does present a challenge to the Gricean approach. The point is as follows.

If I utter something meaning that P, Grice says, then I must intend my hearer to believe that P—or, in a later version, to believe that I believe that P.[1] Against this, Searle says that a speaker's primary purpose is to be understood, and that he may not 'care a hang' whether the hearer accepts either the message or the speaker's sincerity.[2] Grice can allow for understanding without belief—the hearer understands the utterance because he sees what belief the speaker intends him to acquire; but this still represents the speaker as primarily *intending* to produce belief. The question is: what account can we give of *understanding* which will be consistent with the claim that the primary purpose of speakers in general is to produce understanding in their hearers? It is in answering this that Searle finds a special role for intentions about conventions, by equating 'I intend you to understand my utterance' with 'I intend you to recognize the conventional meaning of my utterance'. This does not obviously presuppose that I have any intentions regarding your acquiring of beliefs (other than the one about the conventional meaning), and so Searle's point is met, perhaps.

[1] H. P. Grice, 'Utterer's Meaning, Sentence-Meaning, and Word Meaning', *Foundations of Language* 4 (1968), 225–242, reprinted in Searle (ed.), *The Philosophy of Language*; see p. 230, or pp. 58–59 in Searle's reprint.

[2] Searle, 'Introduction' in *The Philosophy of Language*, p. 10; see also *Speech Acts*, pp. 47–48.

There are a couple of difficulties. (1) The notion of conventional meaning needs to be explained somehow without reinstating the view that an intention to produce belief is after all central to meaning. In Searle's actual analysis, the phrase 'conventional meaning' does not occur, its work being done by 'the rules governing' the utterance; but the problem is there all the same. (2) Would Searle say that in cases of communication which do not m any way involve conventions or rules, the speaker's primary intention is to produce belief? At any rate, the primary intention in such cases cannot be to be understood, if understanding involves grasping rules or conventions. One possibility, for cases where there is no convention, is that the speaker primarily intends the hearer to be led through the Gricean mechanism to *have a certain thought*—perhaps not to believe that he is threatened with a coconut, but to think about being threatened with a coconut and to realize that the speaker intends him to think about this. But then that account might fit also when conventions *are* involved. In short, even if we agree with Searle that a speaker's' primary intention is to be understood, and thus that 'understanding' must be explained in some way other than 'recognizing what belief the speaker intends the hearer to acquire', it is arguable that this result can be secured, quite generally, without referring to rules or conventions.

As for whether Searle is right about the speaker's primary intention: before agreeing with him, I'd like to see some arguments which didn't rely on the parochial and question-begging language of 'illocutionary' and 'perlocutionary'. But all that is by the way. I have accepted a certain conditional of the form [intending]→[meaning], but I have not availed myself of its converse; and so I need take no stand on whether meaning must standardly involve intention to produce belief. In my final section I shall accept two conditionals of the form [meaning]→[intending]: but both are weaker than the one Searle is objecting to—and indeed neither concerns belief.

10. Flouting meaning-conventions

Searle also attacks Grice's intending-to-meaning conditional in another way. Express the conditional as '[intend P]→[mean P]': then what Searle does is to offer a story in which, he says, [intend P] & ¬[mean P]'. I contend that the story fails.[1] Its protagonist is said to intend his hearers to guess that a certain German sentence means that the speaker is a German soldier; but how he can sanely expect them to make this guess, and thus how he can intend them to do so, is not explained. To 'make the case more plausible', Searle suggests that the hearers might be expected to base their guess about the sentence upon their independent belief that the speaker is a German soldier;[2] but the story bears upon Grice only if the speaker intends the hearers to move in precisely the opposite direction, basing their belief that he is a German soldier on their guess about what his sentence means. I cannot see how to make the story both plausible and relevant to Grice's antecedent.

So Searle has not refuted the Gricean conditional [intend P]→[mean P]' because he has not instantiated the antecedent. There is, however, an aspect of Searle's handling of his story that still needs to be discussed.

As an approach to it, consider this remark of Searle's about the 'point' of his story. 'The point of the counter-example', he says, 'is to illustrate the connection between what a speaker means and what the words he utters mean.'

[1] Following Grice, 'Utterer's Meaning and Intentions', at pp. 161–2; and D. M. Armstrong, 'Meaning and Communication', *Philosophical Review* 80 (1971) at pp. 440–441.

[2] Searle, *Speech Acts*, p. 44 n.

It is not clear what this 'connection' is supposed to be, or how the story is supposed to illustrate it. It certainly has nothing to do with the themes of Searle's mentioned in my preceding section. Nor can it be a 'connection' which is involved in Searle's 'revised analysis'; for in that the analysandum is 'S utters sentence T and... means literally what he says', whereas the essential point about the supposedly Grice-refuting story is that in it the speaker does *not* mean literally what he says.

So the 'point' of Searle's story must introduce some fresh theme—and it seems to concern his reasons for saying that in this story the protagonist does not mean that he is a German soldier ('mean that P', for short).

In the first version of the story, the speaker is said not to mean that P 'because what the words mean is' something other than P.[1] Also, of Wittgenstein's challenge to say 'It's cold' and mean 'It's warm', Searle says: 'The reason we are unable to do this is that what we can mean is a function of what we are saying. Meaning is more than a matter of intention it is also a matter of convention.' And he complains that Grice's analysis 'fails to account for the extent to which meaning is a matter of rules or conventions', and does not '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'.

These remarks strongly suggest that a speaker cannot give his utterance anything but its conventional meaning, if it has one. In the later version, however, each remark is toned down.[2] The protagonist fails to mean that P, now, 'because what the words mean and *what he remembers that they mean*' is something other than P. As for Wittgenstein's challenge, Searle now explains why we are 'unable to do this *without further stage-setting*', as follows: 'What we can mean is *at least sometimes* a function of what we are saying. Meaning is more than a matter of intention, it is also *at least sometimes* a matter of convention.' Finally, Grice's analysis is said to leave unexplained 'the extent to which meaning *can* be a matter of rules or conventions', and to leave obscure the fact that what a speaker means is 'more than just *randomly* related' to what his sentence means.

These revised versions are safer, but less clear. Without great confidence, I conjecture that Searle wants to emphasize and explain the fact that a speaker often or usually has no option but to mean by his utterance what he thinks it conventionally means. His remark which I quoted earlier—implying that an utterance's conventional meaning is not 'just another circumstance' helping to determine what the speaker intends (and thus what he means) by it—might have been making the same point: namely that the utterance's conventional meaning is not 'just another' but is a *uniquely powerful* circumstance which in most cases will totally dominate and determine what the speaker can mean by the utterance.

If this interpretation of Searle is wrong, then I do not understand the passages I have recently quoted. If it is right, then one may fairly comment that Searle himself does not explain the facts he wants explained. An analysis of 'saying something and meaning it' cannot throw light on any obstacles there might be to saying something and not meaning it—that is, to meaning by one's utterance something other than what (one thinks) it conventionally means.

Anyway, there is something to be explained in this area. Wittgenstein's remark about saying 'It's cold' and meaning that it's warm does point to a difficulty about wilfully flouting known meaning-conventions. My final task in this paper

[1] All quotations in this paragraph are from 'What is a Speech Act?' pp. 229–230.

[2] All quotations in this paragraph are from *Speech Acts*, pp. 43–5. The italics are all added.

will be to show that the facts about this can be explained, cogently and in depth, on a Gricean basis.

11. A Gricean explanation

Wittgenstein's remark occurs in a context where he is arguing (1) that meaning is not an activity engaged in while uttering, which could be turned on and off at will.[1] Embedded in that is the thesis (2) that one cannot choose what to mean by a given utterance as one could if meaning were an activity synchronous with uttering. And associated with that is the thesis (3) that usually one cannot give to one's utterance anything but its conventional meaning, if it has one. I am to explain (3), but it is best to approach it through (1) and (2).

Take the fairly uncontroversial premise that intending is not an activity engaged in while acting, and add to it the Gricean doctrine that meaning something by what you utter is a special case of intending something by what you do. These jointly imply the Wittgensteinian thesis (1) that meaning something is not an activity engaged in while uttering. By rejuvenating the idea that meaning is a kind of intending, Grice confirms and deepens Wittgenstein's important insight about meaning's not being a process or activity.

To explain the fact (2) that one cannot simply choose what to mean by S, I need to take from Grice a conditional from meaning to intending. It is not the strong, false converse of the intending-to-meaning conditional which I adopted in Section 2. All I need is the extremely weak conditional 'If someone means something by what he utters to a hearer, he intends it to have a certain effect on the hearer', together with the supplementary remark that what he means depends upon what he intends.

Now, to intend to achieve T by a given action one must believe that the action has an appreciable chance of leading to T—the vagueness of 'appreciable' matching the haziness of the line between 'intending to achieve T' and 'acting in the hope that T will ensue'. For a given action X there are countless upshots T such that one cannot perform X intending thereby to achieve T, because the beliefs one has at that time rule out one's having that intention. There could perhaps be someone who did X intending to achieve T, and perhaps one could oneself become able to do X intending to achieve T; but there is, at the given time, an obstacle to one's performing that action with that intention. Furthermore, one cannot remove this obstacle at will, just by choosing to do so; for it consists in the lack of a certain belief, and beliefs are not in that way at the command of the will. And so, by transitivity, one cannot freely choose what to mean by a given utterance. In brief: free choice of meaning requires free choice of intention, which requires free choice of belief, which is impossible.

(As Pascal noticed, one can choose to act in ways which are likely to produce beliefs in oneself, and perhaps we can imagine becoming very and clever in thus manipulating our own beliefs. I don't think that this could count to our being able freely to choose what to believe, but I needn't argue the point. If one could somehow become able freely to choose any given belief, then one could in that way become able freely to choose what to mean by any given utterance. I have implied only that meaning-limits have the same strength as belief-limits, whatever that strength is.)

The foregoing explanation of why we cannot simply choose meanings is due to Donnellan.[2] Satisfactorily, it shows why one cannot simply choose to mean by S that P, yet does not imply that there is any S and P such that one could not

[1] L. Wittgenstein, *Philosophical Investigations*, Oxford 1953, p. 410.

[2] Keith Donnellan, 'Putting Humpty Dumpty Together Again', *Philosophical Review* 77 (1968), 203–215; see especially p. 212.

mean by S that P in some circumstances. That any sentence *could* carry any meaning, if the conditions were right, is essential to the service that convention renders; for what a meaning-convention does is just to make the conditions right for S to mean that P. One might make the general point by saying that for any given S and P it would be possible, after suitable 'stage-setting', to utter S and mean by it that P. I am not suggesting that Searle based his remark about 'stage-setting' on Donnellan's line of thought. Rather, I am complaining that he didn't.

Now we can explain the fact (3) that usually the conventional meaning of an utterance is forced upon the speaker, i.e. is the only thing he can mean by that utterance on that occasion. For this explanation, I take from Grice a stronger conditional than the one I used in explaining fact (2). What I now need is: 'If someone utters something to a hearer, meaning by it that P, then he intends the hearer to associate his utterance in some way with P.' This is still much weaker than the converse of the intending-to-meaning conditional which I adopted in Section 2, for it has both a stronger antecedent and a weaker consequent. I think it is weak enough to be free from serious counterexamples. In addition, I need only invoke the basic fact that in most speaker-hearer situations where someone utters a sentence S,

(a) both speaker and hearer know S's conventional meaning;
(b) neither speaker nor hearer has special reason to associate S with any one proposition other than that which is its conventional meaning;
(c) the facts (a) and (b) are mutually known by speaker and hearer.

Lewis's account of convention enables us to explain why (a), (b) and (c) are usually true; for it implies that it is of the essence of conventions that where they are in force they are mutually known and are in a certain way free from serious rivals. For present purposes, though, it suffices that (a), (b) and (c) are usually true, never mind why. In a situation where they are true, the speaker has a set of beliefs which do not permit him to intend the hearer associate the utterance with anything but its conventional meaning, and so—by my latest meaning-to-intending conditional—the speaker cannot mean by his utterance anything but what it conventionally means. Here, as before, 'cannot' does not mean 'could not in any circumstances'. The speaker is bound to the conventional meaning just to this extent: given that (a) through (c) are true, he cannot mean by S anything except what S conventionally means; and when (a) through (c) are true the speaker cannot make any of them false just by choosing to do so.[1]

[1] In writing this paper I have been much helped by comments, criticisms and suggestions from Michael Beebe, D. G. Brown, S. C. Coval, Keith Donnellan, David Lewis, Howard Jackson, Edwin Levy, and Thomas E. Patton.