Grice and Marty on Expression

Guy Longworth[1]
University of Warwick
December 20th 2015

1. Introduction.

Paul Grice's essay "Meaning" was published in 1957, almost ten years after its initial presentation to a 1948 meeting of the Oxford Philosophical Society. The essay contains an account of what Grice calls "non-natural meaning"—roughly, what speakers mean by, or in, performing certain actions. According to that account, the non-natural meaning of a speaker's utterance—what the speaker meant by, or in, producing the utterance—is determined by a distinctive range of the audience-directed intentions with which the speaker produced the utterance. Grice supported his account by appeal to fruitful, counterexample-based arguments against competitor accounts. Grice's account, and the accompanying defense, has generated a great deal of critical discussion, and now figures in the assumed background to most engagements with its topic.

Important philosophical work rarely forms in a vacuum. Grice cites only one other text in the essay, C. L. Stevenson's (1944) *Ethics and Language*. But that citation connects his essay with a rich body of prior work. Furthermore, Grice's account furnishes a distinction, within the totality of the intentions with which speakers act, between those intentions that constitute acts in which something is meant, or expressed, and those that constitute acts that only serve, or are served by, acts in which something is meant. It thus aims to develop his teacher J. L. Austin's earlier distinction between speakers' illocutionary acts and the other acts that serve, or are served by, those illocutionary acts—specifically, the locutionary acts on which the performance of the illocutionary acts depends, and the perlocutionary acts that depend on the performance of the illocutionary acts. (See Austin 1962, a record of lectures given in 1955. Austin claimed to have formed the views underlying his 1955 in 1939. Related views figured in the background to his 1946.)

[1] I'm grateful for discussion to Laurent Cesalli, Tom Crowther, Naomi Eilan, Hemdat Lerman, Kevin Mulligan, Anne Reboul, Matthew Soteriou, Mark Textor, and audiences in Einsiedeln and London.

In what follows, I discuss a further purported connection between Grice's essay and earlier work. It has been suggested that credit for originating the account offered in Grice's essay is due, not to Grice, but to the Swiss philosopher Anton Marty. For Marty, it has been argued, presented an account of meaning in his 1908 that is similar to, if not identical with, Grice's. (See e.g. Liedtke 1990 and, more tentatively, Cesalli 2013 and Rollinger 2014.) My aim here is to pursue two questions that arise from this priority dispute. To what extent is Grice, as opposed to Marty, responsible for the account presented in Grice's essay? And what is to be learned from pursuit of this particular dispute over priority? My answer to the first question will be that Grice is largely responsible for the account, although it is reasonable to think that Marty's account figures as a sort of precursor. My answer to the second question is that what there is to be learned from pursuit of the priority question has at least as much to do with differences between Marty's and Grice's respective account as it has to do with their similarities. More substantively, I'll discuss a particular difference in emphasis between Grice's account and Marty's, and the way in which it highlights the problematic status of one central component of Grice's discussion.

I'll proceed as follows. §2 ("Precedents and precursors") contains some general reflections on what we might hope to achieve by pursuing historical questions of priority, and some ways in which pursuing answers to such historical questions can interact with the pursuit of questions of a more systematic nature. §3 ("History") considers the question whether the production of Grice's essay could have been influenced by Marty's work. §4 ("Commonalities") explores some similarities and differences between Marty's work and Grice's that are visible at different levels of resolution, and suggests that the central insights present in Grice's essay are absent from Marty's work. §5 ("Expression") considers an apparent problem that Grice presented for the natural idea that intentional behaviour can manifest or express bits of psychology other than the intention with which the behaviour was produced. Marty and Grice offer different responses to this apparent problem: Grice treats it as a genuine problem; Marty treats it as merely apparent. I suggest that one thing that we can learn from Marty—and from historical reflection more generally—is that claims which might otherwise have struck us as uncontestable might be contestable, since contested. I conclude, in §6, by suggesting that we may have as much to learn from differences amongst thinkers in our subject's history as we do from their commonalities.

2. Precedents and precursors.

Before embarking on a more detailed comparison of Grice's account with Marty's, it will be worth pausing to reflect more generally on what we might hope to gain from such an historical investigation.

At the most general level—and without wishing to suggest that these categories are either exclusive or exhaustive—our interests might be more purely systematic or more purely historical. Pursuing the former interest, we might hope that our inquiry will serve more systematic ends, for example by helping us to develop or assess contemporary accounts of meaning. Here, it might be that historical reflection reveals ways in which some aspects of contemporary theories

were originally due to pressures of authority, or more generally that they are contingent accretions or accidents of development, rather than secure bases for further theorizing. In that case, the result of historical reflection might be the weakening, or problematizing, of assumptions that had figured as more or less fixed points in contemporary theorizing. Alternatively, it might be that historical investigation reveals a fund of arguments, or nuances, on which more recent work built without replicating, so that uncovering those arguments and subtleties can figure in strengthening contemporary positions. Pursuing the latter interest, we might hope that our inquiry will serve more historical ends, for example by deepening our understanding of the crisscrossing paths along which contemporary accounts have developed.

Insofar as we adopt a more purely historical view, we are liable to do so because we are concerned with the path along which what we take to be insights have developed. Thus, we are liable to be interested in the historical development of what we take to be contemporary successes. Here it's worthwhile to distinguish two ways in which earlier work might be related to later work. The first way is for the earlier work to be a *precedent* for later work: minimally, for it to involve an earlier appearance, in whole or part, of later insights. Here, what matters are commonalities or similarities between the earlier and later view, whatever the historical connections between those views. The second way is for the earlier work to be a *precursor* for the later work: minimally, for the earlier work to have figured causally in the generation of the later work, either as a prompt to, or stage in, the development of the insights contained in the later work. In principle, one might find perfect similarities between earlier and later views without the former serving as precursor for the latter: that is, one might find mere precedence. Similarly, one might in principle find precursors for later views with little similarity to those later views, for example if the earlier views had served merely as prompts, or argumentative foils, in the development of the later views.

The assessment of claims about precedence, and so of claims about commonality or similarity, is liable to be sensitive to the values of two interrelated parameters. First, it is liable to be sensitive to the level of resolution at which commonality, or similarity, is sought. Second, it is liable to be sensitive to what we take to be the central insights of later work. Thus, views that appear very similar at a low level of resolution may be strikingly dissimilar at higher levels of resolution, and the key insights of later views might be discernible only at those higher levels of resolution. For example, we might take the insights of some later views to be found not so much in the accounts that they offer, but rather in the specific arguments that they present in defense of those accounts. The interplay between questions about precedence and questions about the level of resolution at which a view's central insights are best discerned provides another point at which systematic and historical interests intersect.

For example, at one level of resolution, we might consider Aristarchus of Samos's early heliocentric model of the Universe to be a precedent for Copernicus's model of the solar system, since the accounts have in common that the Earth is taken to orbit the Sun. However, at a higher level of resolution—a level that makes more clearly visible what we take to be Copernicus's central insights—we would attend to the very different ways in which the two thinkers

sought to defend their models: the former, through crude calculations of the relative sizes of the Earth and the Sun conjoined with an appeal to the implausibility of the larger orbiting the smaller; the latter, through the relative simplicity of his more detailed heliocentric model of the solar system. Focusing on arguments, rather than broad claims, we would be less inclined to treat the earlier model as a precedent for the latter. (We might nonetheless treat Aristarchus's account as a precedent of sorts for the basic heliocentrist claim. And Copernicus himself may have treated Aristarchus's claim as a partial precursor for his own.) (See e.g. Gingerich 1985.)

With those general reflections in hand, let's consider how they apply to our specific target, the priority dispute between Marty and Grice. We'll begin with the more historical question, whether, or to what extent, Marty's account is causally connected with Grice's (§3), before turning to the question of the extent to which there are important commonalities between Marty's account and Grice's (§4).

3. History.

What are the grounds for thinking that Grice's discussion in "Meaning" could have drawn content, or inspiration, from Marty's earlier work? Had Grice read Marty's work? If not, had he read those who had read, and perhaps cited, Marty work? As far as I'm aware, there are no clear grounds for thinking that Grice had read Marty's work before composing "Meaning". However, there may yet be reason to suppose that Grice had indirect access to some of the content of that work. At a recent conference on Marty's work, Kevin Mulligan briefly sketched a positive case for thinking that Grice had such access. I want here briefly to record the basis of Mulligan's case, as it struck me. Mulligan should be held responsible for whatever truth there is in the account offered here; for the rest, responsibility is mine.

Mulligan's proposal takes off from the central role played in early 20th Century thinking about language by *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism* (1923) by C. K. Ogden and I. A. Richards. Ogden and Richards' book had an enormous impact on philosophy and linguistics from its publication until at least the mid-50s. Although Grice doesn't cite the book, it is very likely that he had read it. (Grice tends against full citation, though does a reasonable job of mentioning individual influences. However, more evidence would be needed before we could be certain that Grice read, or read carefully, this text.) We might add here that Grice had certainly read Stevenson 1944, since that work is cited in "Meaning". And Stevenson both cited, and actively engaged with, Ogden and Richards' text in the chapter that Grice cited. Let's assume that Grice had read their book, and that it had exerted some form of influence on his discussion in "Meaning". The question now is whether Ogden and Richards' book could have served as a conduit for Marty's work.

Initial inspection is disappointing. Ogden and Richards include no reference to Marty in their index. However, as Mulligan noted, they did discuss other potentially relevant work. In particular, they discussed work by the archaeologist Alan Gardiner, quoting sympathetically the following passage:

> Is the meaning of a sentence that which is in the mind of the speaker at the moment of utterance or that which is in the mind of the listener at the moment of audition? Neither, I think. Certainly not that which is in the mind of the listener, for he may utterly misconstrue the speaker's purpose. But also not that which is in the mind of the speaker, for he may intentionally veil in his utterance the thoughts which are in his brain, and this, of course, he could not do if the meaning of the utterance were precisely that which he held in his brain. I think the following formulation will meet the case: The meaning of any sentence is what the speaker intends to be understood from it by the listener. (Gardiner, 1922: 361)

As we'll see in the next section, Gardiner's proposal is strikingly similar to—though not identical with—Grice's account, and is taken very seriously by Ogden and Richards. That is of some independent interest. But does it aid our search for a connection back to Marty? Specifically, can we trace a line of influence from Marty to Gardiner?

We can. Gardiner cites Marty (1908) in a very positive way:

> Most writers on Languge have, of course, been more or less alive to this standpoint [—the standpoint of attending to speakers' purposes in theorizing about language, GL], but Marty alone, so far as my reading goes, is entirely impregnated with it. His statement of the purpose of Language agrees closely with my own definition, which runs: Language is the name given to any system of articulate symbols having reference to the facts of experience, whereby speakers seek to influence the minds of listeners in given directions. (Gardiner, 1922: 354.)

So, Gardiner provides us with the remaining portion of our path back to Marty. Disappointingly, as the quote indicates, Gardiner finds in Marty a more or less generic appeal to speakers' intentions and, more specifically, their intentions to influence listeners' minds. As we'll see in more detail in the following section, that means that Gardiner's own proposal is closer to Grice's than is the less specific view that Gardiner admits to finding in Marty. Insofar, then, as the path of possible influence that we have traced from Marty to Grice—via Gardiner, Ogden and Richards, and Stevenson—is the only, or main, line of influence, the most that we can say is this. Marty's work, as so transmitted, may have suggested to Grice, or supported his standing interest in, the idea that speakers' intentions to influence the minds of listeners are important determinants of what speakers mean. However, as we will see in the following section, the specificity of Grice's positive account, together with his objections to other views that also attend in this general way to speakers' intentions, means that the connection sketched here is less decisive than some enthusiasts might have hoped. Nonetheless, it seems reasonable, in light of Mulligan's proposal, to hold that Marty's work exerted at least an indirect influence on Grice's. In the following section, I turn to considering the more delicate question of the precise shape that that influence took, or could have taken.

4. Commonalities.

In "Meaning", Grice presents his account in the following way:

> *A* meant$_{NN}$ 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 may add that to ask what *A* meant is to ask for a specification of the intended effect (Grice 1957: 220)

Grice presents this as an account of what he calls non-natural meaning, $\text{meaning}_{\text{NN}}$. (The contrast, to which we'll return, is with cases of natural meaning, including the connection between smoke and fire when the smoke means fire.) As noted, the question of precedence depends on the prior, systematic question of what we take to be Grice's central insight or insights. As Grice presents his account, we might focus on any of the following candidate insights (presented at increasingly high levels of resolution):

1. The claim that what a speaker means by an utterance is determined by their intentions.
2. The claim that what a speaker means by an utterance is determined by their intentions to have a specific range of effects on an audience's psychology.
3. The claim that what a speaker means by an utterance is determined by their intentions to have a specific range of effects on an audience's psychology by means of the audience's recognizing certain features of the speaker's psychology.
4. The claim that what a speaker means by an utterance is determined by their intentions to have a specific range of effects on an audience's psychology by means of the audience's recognizing the speaker's intention to have certain effects on the speaker's psychology.
5. The claim that what a speaker means by an utterance is determined by their intentions to have a specific range of effects on an audience's psychology by means of the audience's recognizing that very intention—that is, the claim that what a speaker means is determined by some of their specifically reflexive intentions.

The five insights differ in specificity and, in particular, with respect to the specification they offer of the intentions that are taken to determine what speakers mean. The level one insight is just that a speaker's intentions figure in determining what the speaker means. The level two insight entails the level one insight, but goes beyond it in requiring that it is a speaker's intentions specifically to influence their audience's psychology that determines what they mean. The insight at level three entails that at level two but goes beyond it in requiring that the operative intention is to influence the audiences' psychology by means of the audience's recognition of something about the speakers' psychologies. The level four insight outstrips the level three insight by requiring that influence on the audience's psychology is to be achieved through the audience's recognition of the speaker's intention to have such effects. Finally, the level five insight entails the level four insight by embodying that insight in specifically *reflexive* form: the intention taken to determine what the speaker means is an intention to influence the audience's psychology by means of the audience's recognition of that very intention. It is the insight at level five that most closely tracks Grice's explicit presentation of his position in the passage with which we began.

The question of influence depends, then, on three questions. First, at which level of resolution should we locate Marty's central insight(s)? Second, at which level of resolution should we locate Grice's central insight(s)? Third, do the insights visible at those levels of resolution match: is Grice's central insight reasonably viewed as a version of Marty's, or do their respective insights occur at different levels of resolution? To forestall suspense, I'll be suggesting that Marty's

central insight is visible at around level three, whilst Grice's insight is most clearly seen at levels four or five. So, although there are clear similarities between Marty's position and Grice's at lower levels of resolution, Grice's central insight isn't yet visible at those levels of resolution. Put another way, although some of Marty's insights are reflected in Grice's work, Grice's central insights are not to be found in Marty's. I'll begin to develop a case for that verdict by considering the first question: what was Marty's insight? Let's consider some passages in Marty's work to which appeal has been made in attributing an intention-based account of meaning to Marty.

Marty provides an initial summary of his position in the following passage:

> We said that language…is primarily understood as the intentional indication of inner life. However, what is primarily intended in this indication is a corresponding influence of foreign inner life. As a rule, one indicates one's own presentations, judgments, emotions &c., in order to trigger presentations, judgments and emotions in another psychical being, and indeed, ones which are analogous to one's own. (Marty, 1908: 22, cited by Cesalli: 144.)

Here, Marty's proposal seems to be the following. Language use is a form of intentional activity: one speaks in order to attain certain ends. More specifically, Marty suggests two intentions that might figure in shaping one's linguistic activity. First, in speaking one intends to "indicate" specific aspects of one's own inner life. Thus, for example, in speaking to an audience, one might aim to indicate, or to make recognisable, that one harbours, say, a specific belief or feeling. However, Marty suggests that that would be a secondary aim, a means to one's primary end. Second, then, one's primary aim in indicating to one's audience this aspect of one's inner life would be to exert a "corresponding influence" on their inner life. That generic characterisation leaves open whether the primary intention would be served merely by the audience's recognition that one harbours the indicated belief, feeling, or other aspect of inner life. That would be one way in which one's indication might figure in exerting a corresponding influence on one's audience's inner life. However, Marty appears to have something more specific in mind—namely, that one's primary intention would be that one's audience not only come to recognise that one harbours the indicated aspect of inner life, but that they should thereby come to harbour an "analogous" aspect of inner life. In cases in which the indicated aspect of inner life is a belief, it would be natural to view the aim as being that the audience should come to have the same, or approximately the same, kind of belief. In cases in which the indicated aspect of inner life is a feeling, that would be less natural. Perhaps the proposal in those cases would be that the audience needn't come to share the same kind of feeling, but that they would nonetheless need to go beyond mere recognition of the feeling, perhaps by engaging in some form of empathy.

Marty doesn't here quite commit to the claim that either or both of the intentions that he mentions figures in determining what the speaker means by their utterance. But that would be a natural extrapolation, and is something to which Marty appears elsewhere to commit. When so construed, we can certainly find in Marty the level one insight: the claim that what a speaker means by an utterance is determined by their intentions. Furthermore, we can also find the level two insight: the intentions to which Marty appeals are intentions to have a specific range of effects on an audience's psychology, specifically that they come

to recognise and thereby to share (or e.g. appropriately to empathize with) a target aspect of the audience's inner life. Finally, we can reach the level three insight: focusing on the primary intention as the central determinant of what the speaker means, we can see that that overarching end is to be attained via the audience's recognition of the aspect of inner life that the speaker indicates. What we don't find is the more specific level four insight, according to which the primary intention is to be achieved via the audiences recognition specifically of one or other of the speaker's intentions, as opposed to other aspects of their inner life. Although that insight is consistent with what Marty says, it nonetheless goes beyond anything that he makes explicit. (See Cesalli 2013: 143 fn.18 and Liedtke 1990: 43–44, both of whom seem to recognise analogues of this point.)

Compare, now, a later passage in which Marty's intention-based account of meaning appears:

> The name signifies the idea, the statement signifies a judgment, and what we call the meaning of an expression is in each case that very content of the soul to evoke which in the hearer is its essential reference and its ultimate aim (be it by nature or by habit), provided that at the same time it has the ability to reach this aim as a rule. Nevertheless, it never reaches this aim immediately, but only by being at the same time a sign of the psychological phenomena in the speaker. (Marty 1916: 68f., cited by Liedtke 1990: 30.)

Here, Marty makes explicit that the meaning of an expression (as opposed, perhaps, to what a speaker means by an utterance of that expression) is determined by "ultimate aim"—so, plausibly, by intention. Specifically, meaning is identified with the aspect of inner life that the speaker intends the audience to come (approximately) to share. As before, that ultimate end is to be achieved by means of the audience's recognition of corresponding, or analogous, psychological phenomena in the speaker. But, again, there is no explicit commitment to the level four insight that the operative psychological phenomena are the speaker's intentions. As far as I've been able to ascertain, Marty doesn't anywhere make claims about meaning that would compel us to locate his position at level four or above. If that is right, then the question whether Marty has priority turns on the question whether Grice's central insight is visible only at levels four or five.

Before turning to the question where Grice should be located on our list of insights, it will be worth considering briefly where Gardiner should be placed. Recall his formulation of his position—and, by extension, the position that he attributed to Marty's: "The meaning of any sentence is what the speaker intends to be understood from it by the listener." (Gardiner, 1922: 361) Gardiner, it seems, doesn't go further that level two. That is, he claims only that what a speaker means by an utterance is determined by their intentions to have a specific range of effects on an audience's psychology—in particular, their intention that the audience should comes to understand an uttered sentence or utterance. Gardiner says nothing here about the role of the audience's recognition of aspects of the speaker's psychology in facilitating their understanding. If Gardiner served as Grice's only conduit, it is therefore plausible that his access to Marty's work would have been confined to Marty's level two insights.

Let's turn, then, to the question of where Grice's position should be located on our list of insights. We'll focus here more or less on Grice's own view about

that. Grice certainly sought to go beyond level two, since he argued explicitly against a position that rested at that point:

> A first shot would be to suggest that "*x* meant_{NN} something" would be true if *x* was intended by its utterer to induce a belief in some "audience" and that to say what the belief was would be to say what *x* meant_{NN}. This will clearly not do. I might leave *B*'s handkerchief near the scene of a murder in order to induce the detective to believe that *B* was the murderer; but we should not want to say that the handkerchief (or my leaving it there) meant_{NN} anything or that I had meant_{NN} by leaving it that *B* was the murderer. (Grice 1957: 217.)

Although Grice argued for the need to go beyond level two, he leaped immediately to endorsement of a level four insight, without pausing to consider the standing of the less specific insight available at level three:

> Clearly we must add at least that, for *x* to have meant_{NN} anything, not merely must it have been "uttered" with the intention of inducing a certain belief but also the utterer must have intended an "audience" to recognize the intention behind the utterance. (Grice 1957: 217.)

However, it is plausible that Grice would have taken an argument to be available against remaining at level three that is analogous to the argument that he brings against remaining at level two. Thus, for example, I might try to arrange things so that an audience could come to see, on the basis of observing some behaviour of mine, that I believe that *B* is the murderer, with the intention of inducing the audience to come also to believe that *B* is the murderer. As in Grice's handkerchief case, it seems implausible to suppose that this would suffice for my meaning_{NN} anything by my belief-revealing behaviour. Again, a minimal fix would be to add a requirement that there be an intention on my part that, on the basis of observing my performance, my audience is, in addition, to recognize my intention that they come to believe that *B* is the murderer.

Grice focuses here on the insufficiency for meaning_{NN} of a speaker's intending to influence the audience's psychology by means of the audience's recognition of a piece of the speaker's psychology. However, it's worth noticing that an analogous point could be made with respect to the claim that this is a necessary condition. For although it is plausibly a necessary condition on meaning_{NN} that, as Grice claims, the speaker should intend the audience to recognize their intention, it is not plausibly a necessary condition that the speaker should intend the audience to recognize anything else about their psychology. For instance, it seems obvious that a speaker might mean by an utterance that *B* is the murderer without the speaker in fact believing that *B* is the murderer. Now that seems to be so even if the speaker should know that they do not believe that *B* is the murderer. And it seems to be so even if, in addition, the speaker knows that one cannot recognize what isn't so. Since one cannot reasonably intend what one knows to be impossible, a speaker in that position cannot reasonably intend their audience to recognize that they believe that *B* is the murderer, and so cannot reasonably intend to influence the audience's psychology by means of the audience's recognizing that. However, none of that seems to stand in the way either of the speaker meaning something by what they say or—connectedly, according to Grice's level four insight—of their intending the audience to recognize their intention to influence the audience to believe that *B* is the murderer.

If that line of thought is correct, then there is reason to hold, not only that Grice thought that there were good reasons to go beyond level three, but also that Grice was correct to think that. In that case, it seems reasonable to identify Grice's central insight as occurring either at level four or at level five. Thus, it seems reasonable to view Grice's insight as going beyond anything to be found in Marty's work. There are commonalities between Grice's work and Marty's. However, significant differences emerge at higher levels of resolution. Grice's contribution to our understanding of speaker meaning plausibly goes beyond the commonalities. It would therefore be misleading to treat Marty as having priority except with respect to the lesser insights that the two thinkers share.

Although the commonalities between Grice's work and Marty's are comparatively uninteresting, there may yet be something to learn from some of their differences. One such difference is the one recently isolated. Marty thinks that an audience's recognition of elements of a speaker's psychology other than their intentions can play a central role in communicative transactions, whilst Grice thinks that only recognition of speaker's intentions figures in determining meaning. When presented in that way, it is possible to endorse both viewpoints, since it would be consistent with denying that elements of a speaker's psychology figure in determining what the speaker means to accept that the elements play other important roles in communication. However, in the following section we'll consider a way in which Grice's emphasis on intention over other elements of psychology induces an apparent problem. And we'll see that Marty's counter-emphasis might provide the resources for a satisfying dissolution.

5. Expression.

Grice reflected further on the topic of non-natural or meaning in "Meaning Revisited," written around twenty years after "Meaning" was published. (1982, composed around 1976.) In the later essay, Grice attempted to provide a more synoptic perspective on the relation between natural and non-natural meaning, and on his account of non-natural meaning.

Grice's discussion begins by emphasising a way in which natural meaning can figure in revealing aspects of psychology:

> In the case of natural meaning, among the things which have natural meaning, besides black clouds, spots on the face, and symptoms of this or that disease, are certainly forms of behavior: things like groans, screeches, and so on, which mean, or normally mean, that someone or something is in pain or some other state. (1982: 292.)

Grice's aim is to try to explain ways in which such natural expressions of human psychology as groans and screeches might conceivably have come to be replaced by non-natural expressions as characterised in his account of non-natural meaning. To a good first approximation, the aim is to account for a conceivable transition from *behaviour* that expresses bits of psychology to *subjects* who express bits of their psychologies—for example, the transition from the expression of suffering in a groan to the (intentional) expression of suffering by its subject.

Although Grice doesn't highlight this, it is important to notice that the account of non-natural meaning arguably retains an essential role for behaviour that naturally means bits of its producer's psychology. For the account of non-

natural meaning relies upon the possibility that a producer's behaviour can reveal those of their intentions that figure in determining what they non-naturally mean by their utterances. And the connection between behaviour and intention on which the possibility of such revelation of intention in behaviour depends must be a form of the relation of natural meaning.

Grice's account begins, then, with the idea of pieces of behaviour as naturally meaning bits of psychology, and so as potentially revealing to the appropriately sensitive those bits of psychology. On that basis, Grice seeks to develop, in a series of stages, a conceivable developmental trajectory leading to the birth of non-natural meaning.

> Stage one…involves the supposition that the creature actually voluntarily produces a certain sort of behavior which is such that its nonvoluntary production would be evidence that the creature is, let us say, in pain. The kinds of cases of this which come most obviously to mind will be cases of faking or deception. (1982: 292.)

Thus, we have the idea of a sort of behaviour that—at stage zero—naturally meant, and so potentially provided evidence for, a creature's being in pain. However, at stage one, the creature has acquired the capacity to produce behaviour of that sort voluntarily—that is, as constituting a piece of intentional action. A sort of behaviour that had naturally meant pain now fails to naturally mean pain, at least on those occasions on which the behaviour is the immediate upshot of intention rather than felt pain. At stage one, it is allowed that other creatures do not, and perhaps cannot, keep track of the shift in sources of the target sort of behaviour. Thus, at stage one, it is plausible that those other creatures that were appropriately sensitive to the target sort of behaviour at stage zero will continue to treat voluntary productions of that behaviour as naturally meaning, and so as revelatory of, pain.

The other creatures become cognizant of the voluntariness of some instances of the sort of behaviour at stage two:

> In stage two not only does creature X produce this behavior voluntarily instead of nonvoluntarily, as in the primitive state [stage one], but we also assume that it is *recognized* by another creature Y, involved with X in some transaction, as being the voluntary production of a certain form of behavior the nonvoluntary production of which evidences, say, pain…. The import of the recognition by Y that the production is voluntary undermines, of course, any tendency on the part of Y to come to the conclusion that creature X is in pain. (1982: 293.)

Other creatures will no longer treat as revelatory of pain those instances of the target sort of behaviour that they recognize to be the immediate upshot of intention, rather than of felt pain. Plausibly, that is because they will treat those instances of behaviour as revelatory of the intention with which the behaviour was undertaken and take that to be incompatible with the behaviour also being the immediate upshot of pain.

It's worth noticing that Grice doesn't suggest grounds here for thinking that coming to recognize that some instances of the target sort of pain behaviour have an intentional source would throw into doubt the standing of those instances of such behaviour that do have an immediate source in felt pain. That is, he doesn't seek to present a broader sceptical worry about creatures capacities ever to recognize that a piece of behaviour naturally means pain that might be thought to arise from the fact that a single sort of behaviour is now recruited as an expression

of either pain or intention. And raising that concern would be problematic in this context, since it would seem equally to threaten creatures' capacities to recognize that some instances of the target behaviour naturally mean intention. The specific concern is just that recognizing a piece of behaviour as having its immediate source in intention would render unreasonable treating that instance of behaviour as also having its immediate source in felt pain.

At least by stage two, then, we are presented with an apparent problem. Suppose that one hopes to reveal to another creature that one is in pain, or has a specific opinion, or the like. One might think that a natural way to proceed would be by undertaking behaviour that one takes to be a recognizable expression of pain, opinion, or the like. The problem one now faces is that if one's audience recognizes one's performance as the immediate expression of one's intention so to behave, then it will be unreasonable for them also to treat one's performance as expressing anything other than the intention with which one so behaved.

In recent work, Richard Moran has pressed a version of this concern:

> …if we are considering speech as evidence, we will have eventually to face the question of how recognition of its intentional character could ever *enhance* rather than detract from its epistemic value for an audience. Ordinarily, if I confront something as evidence…and then learn that it was left there deliberately…this will only discredit it as evidence in my eyes. It won't seem *better* evidence, or even just as good, but instead like something fraudulent, or tainted evidence. (Moran 2005: 6.)

As presented by Grice and Moran, the problem emerges only at stage two, the stage at which other creatures come to recognize that a sort of behaviour that had once naturally meant (only) pain has been recruited so as to naturally mean (also) an intention with which the behaviour was undertaken. However, it is arguable that the basic problem emerges earlier, at stage one. For at stage one, we have that a sort of behaviour each instance of which once naturally meant (say) pain now has instances that naturally mean intention. As we noted, the problem presented by Grice and Moran is not the generic sceptical problem that particular instances of that sort of behaviour now mean either pain or intention, and so fail either to mean pain or to mean intention. Rather, it is a more specific problem. We are to suppose that some instances of the sort of behaviour naturally mean intention. The problem is that that intention itself might have been the upshot either of pain or of some other bits of psychology. Thus, the behaviour, in naturally meaning, and so potentially revealing, the intention with which it was undertaken at the same time screens off from immediate revelation in behaviour the pain that one might have hoped to reveal. At best, the pain is the immediate source of one's intention, rather than one's behaviour, and so is not immediately revealed in one's behaviour.

The problem, then, does not distinctively affect only those creatures that recognize that a piece of behaviour has its immediate source in intention rather than pain. Creatures that haven't yet recognized this are no better placed to exploit the natural meaning of intentional behaviour in order to recognize pain, since such behaviour does not in fact naturally mean pain, but (at least in the first instance) only intention. We might allow that those who treat the behaviour as revealing pain whilst failing to recognize the behaviour's intentional source are in some sense more reasonable than those who seek to reconcile treating the

behaviour as revealing pain whilst recognizing that it reveals only intention. But if what Grice and Moran present is a genuine problem, then even at stage one, neither party is in a position to recognize the behaviour as having its immediate source in pain, since one can recognize only that which is so, and the behaviour originates not in pain, but rather in intention.

In the face of the apparent problem, Grice asks,

> …what would be required to restore the situation: what could be added which would be an antidote, so to speak, to the dissolution on the part of [the audience] of the idea that [the producer of the behaviour] is in pain? (Grice 1982: 293.)

In setting up the problem, Grice in effect assumes that where a piece of behaviour is produced voluntarily—as expressive of an intention so to behave—the piece of behaviour naturally means the intention. In the first instance, then, the behaviour fails to be revelatory of any aspect of the producer's psychology distinct from the intention that is immediately responsible for the behaviour. Thus, any intention to behave in a way designed to reveal more than the intention immediately responsible for the behaviour seems bound to be unsatisfied. And cognizance of that fact would make it impossible for a rational subject to form an intention to reveal anything beyond the intention with which the behaviour was undertaken.

Grice's response to the problem has two components. The first component exploits the fact that a piece of behaviour can naturally mean, and so reveal, the intention with which it was undertaken. If that intention is to be satisfiable, it cannot be an intention to reveal in behaviour anything other than itself, for only the intention itself is immediately reflected in the behaviour that it shapes. Thus, Grice ensures that the intention is satisfiable by making it the reflexive intention that it—the intention with which the behaviour was undertaken—be recognisable. (That is, Grice's response to the present difficulty converges with the level five insight characterised above.) In place of the sort of intention involved at stage one, merely to present a sort of behaviour that other creatures would be unable to distinguish from that revelatory of pain, the reflexive intention is to present behaviour that reveals itself as an open revelation of some of one's motives in so behaving.

The second component of Grice's response then seeks to restore a route from such behaviour, via the reflexive intention with which it was produced, to the psychological facts that the performance was designed to make available:

> Whether or not in these circumstances [the audience] will not merely recognise that [the producer] intends, in a certain queer way, to get [the audience] to believe that [the producer] is in pain, whether [the audience] not only recognizes this but actually goes on to believe that [the producer] is in pain, would presumably depend on a further set of conditions which can be summed up under the general heading that [the audience] should regard [the producer] as trustworthy in one or another of perhaps a variety of ways. (Grice 1982: 294.)

In short, what is required would be akin to an inferential transition, via premises to the effect that the producer would not intend openly to communicate that they were in pain (or in some other psychological condition) unless they were in fact in pain (or in the other psychological condition).

One natural worry at this point concerns the audience's entitlement to regard the producer as, in the required ways, trustworthy. The audience's merely assuming this might serve to underwrite their forming the requisite belief that the

producer is in pain, and to that extent serve some of the producer's purposes. But it is implausible that a mere assumption of trustworthiness could in general figure in the audience's acquiring *knowledge* that the producer is in pain. Thus, insofar as the audience or producer cares about whether the audience comes to know that the producer is in pain—as opposed, say, to merely assuming or surmising it—it seems that a further story must be told about their entitlement to regard the producer as trustworthy. And insofar as the specific intention with which the producer acts is taken to screen off any of their further motives from the audience, it is not straightforward to see what account is to be given of the audience's warrant for regarding them as trustworthy. Alternatively, it would need to be explained how merely regarding the producer as appropriately trustworthy—that is, without possessing any warrant for so regarding them—could sponsor a route to knowledge, as opposed to mere opinion, about their psychological condition.

Whatever the fate of such approaches to restoring a route from intentional behaviour to underlying psychology, one might reasonably wonder whether those approaches are needed. For one might reasonably wonder whether the apparent problem that Grice and Moran present is genuine. It is here that an apparent difference between Marty and Grice assumes importance. Marty, unlike Grice, can be read as holding that a piece of behaviour might be both intentional and nonetheless capable of revealing to an audience bits of the producer's psychology other than the intention with which the behaviour was produced.

Marty discusses the way in which behaviour can mean, and so reveal, aspects of the producer's psychology in the following passage:

> Where we have to do with an unintentional indication of our inner life, for example, with an unintentional shout, … [in] such a case, to mean and to be a sign means only to indicate something and to make it known, just as one says of thunder that it is a sign of electric discharge, and of dark clouds that they are a sign of rain…. Just like an involuntary shout, the voluntary uttering of a name or of a statement indicates a piece of the speaker's psychical life. (Marty, 1908: 280–4)

Marty's discussion prefigures one theme in Grice: the idea that a piece of unintentional behaviour can naturally mean, and so reveal, aspects of the producer's inner life. However, Marty seems to diverge from Grice in his conception of the revelatory power of intentional behaviour. For he seems to suggest that the voluntary uttering of a name or a statement, just like an involuntary shout, can naturally mean, and so reveal, aspects of the producer's inner life.

Strictly, of course, that would be consistent with Grice's view, according to which such behaviour can naturally mean, and so reveal, only one sort of aspect of the producer's inner life: the intentions with which the behaviour was produced. However, we've already seen that Marty's account incorporates a broader conception than Grice's of the revelatory power of human behaviour. In effect, it was this broader conception that figured in Marty's failure to attain Grice's specific insights into the nature of non-natural meaning—that is, to attain insights at level four and five.

Laurent Cesalli affirms the proposed reading of Marty in the following passage:

> The fact that [the utterance "Socrates exists" when uttered] efficiently indicates [the [utterer's] judgement: Socrates exists] is independent of [the utterer's

secondary intention, their intention to indicate the/their judgment: Socrates exists], just as thunder indicates lightning independently of anybody's will or intention. (Cesalli, 2013: 147, with interpolations from 145–147)

On this reading, Marty's proposal is that the relation between a piece of behaviour—in this case, an utterance—and a judgement—so an aspect of psychology distinct from the intention implemented in the behaviour—can be akin to the connection between thunder and lightning. That is, the relation can be one of natural meaning. Crucially, Cesalli reads Marty as endorsing that view of the relation between behaviour and judgement even in cases in which the producer intends to indicate, and so to reveal, their judgement to an audience.

Even if we assume that Marty diverges from Grice on this point, the best explanation for that divergence may be simply that Marty has failed to reflect on the potential effects of intention in screening off behaviour from other aspects of inner life. That is, it may be that Marty failed adequately to consider the impact on his favoured model of the relation between behaviour and its psychological sources of allowing that some of its immediate sources are intentions. However, a more tantalising possibility is that Marty's disagreement with Grice on this point is principled, and due to Marty's perception, or seeming perception, of a way in which behaviour can be both intentional and at the same time revelatory of aspects of psychology other than intention.

As far as I'm aware, Marty didn't offer an explicit account of how behaviour that was intentional might nonetheless naturally mean aspects of psychology other than intention. His main contribution at this point arises from his willingness to allow for that possibility. For that willingness reveals to us that there is a question here, one that might have been concealed by Grice and Moran's presentation of their purported problem. In revealing the existence of a question at this point, Marty compels us to address it, rather than simply assuming the correctness of the answer offered by Grice and Moran.

The question whether the purported problem is genuine is large and delicate, and not one that I can address in detail here. I want to conclude this section by suggesting two ways in which an attempt might be made to defend the type of position suggested in the passage from Marty.

The first way in which an attempt might be made to defend Marty's suggestion in effect exploits a way in which he needn't be regarded as in direct conflict with Grice. Grice is committed to the view that his problem arises in cases in which behaviour is intentional in the strong sense of being the immediate upshot of, and so shaped by, an intention so to behave. By contrast, it might be possible to read Marty as contemplating a more generous conception of voluntary behaviour, on which the behaviour is not itself the immediate upshot of an intention, but rather is merely condoned, as in accord with the producer's intentions. The model here would be one on which the behaviour is the immediate upshot of, and so naturally means, a judgement, rather than an intention. It is voluntary, and so intentional, only in the sense that the producer could have intervened by preventing the behaviour and yet didn't, and in that sense condoned the behavioural expression of the judgement. The fact that a piece of behaviour is in that thin sense voluntary or intentional need be in no conflict with its being an immediate expression of aspects of psychology other

than intention. Thus, if Marty wished only to endorse that possibility, it would be possible for he and Grice to both be right about the different cases that they treat.

The first way in which an attempt might be made to defend Marty's proposal is, for at least two reasons, meagre. The first reason is that it leaves open that Grice's problem is genuine with respect to all cases of fully intentional behaviour. The second, related reason is that it would be natural to think that the most interesting class of cases, including most or all cases of voluntary utterance, will be intentional in the more demanding way exploited in generating Grice's problem. That is, it would be natural to think that the activity involved in voluntary speech is the immediate upshot of intention, rather than being an immediate upshot of some other aspect of psychology that could have been, and wasn't, intentionally prevented. Even accepting the possibility of cases of the sort to which appeal was made in the first attempt at defence, one might reasonably hope for a more direct response.

The second way in which an attempt might be made to defend Marty's proposal builds on materials in the first, but constitutes a more direct response. It retains from the first attempt the idea that bits of behaviour can naturally mean, and so reveal, bits of psychology. In order for the behaviour to have that natural connection with the psychology, the idea was, the behaviour must be an immediate upshot of the revealed aspects of psychology. Grice's purported problem is then the following. Where an activity is intentional, it must be the immediate upshot of the intention with which it was undertaken, rather than being the immediate upshot of any piece of psychology distinct from that intention. In that case, it is difficult to see how an intentional activity can naturally mean any piece of psychology distinct from the operative intention. However, in setting up Grice's problem, we followed Grice in assuming that the intention and the competing aspect of psychology have a single sort of upshot, the occurrence of instances of a single sort of behaviour. Given that assumption, we would have a single type of activity—the target sort of behaviour—over the dominion of which intention and other aspects of psychology are forced to compete. The second attempt concedes that if we were compelled to accept that assumption, then Grice's problem would be genuine, but proposes that we need not accept the assumption.

The proposal, then, is that we should reject the assumption that exactly the same upshot must be involved in the target cases of intentional and unintentional behaviour. Rather, we should allow that, in at least some cases, the activity that is the immediate upshot of intention isn't simply the behavioural surface of the unintentional behaviour, but is rather an activity involving the manifestation of psychology in behaviour. In such cases, what the producer intends, and is sometimes capable of achieving, is not simply the replication of a piece of behaviour, but is instead an action that involves the revelation of psychology in behaviour. The behavioural surface is the immediate upshot of aspects of psychology other than intention, and so there is no competition for dominion there. The expressive episode as a whole is the immediate upshot of intention rather than other aspects of psychology, and so there is no competition there either. The other aspects of psychology figure as partly constituting, rather than as causing, the expressive episode by virtue of their role in causing the behavioural

surface of that episode. Their figuring in that way is therefore perfectly compatible with the intention playing the complementary role of causing the expressive episode as a whole, rather than its behavioural surface.

More work would be required in articulating the line of defence all too briefly sketched here before we would be in a position properly to assess it. However, I hope to have said enough to indicate that the question that Marty forces on us has no simple resolution.

6. Conclusion

Let me recap. §2 ("Precedents and precursors") contained some general reflections on what we might hope to achieve by pursuing historical questions of priority, and some ways in which the answers to such historical questions can interact with the pursuit of questions of a more systematic nature. §3 ("History") considered the question whether the production of Grice's essay could have been influenced by Marty's work, and drew on a suggestion of Kevin Mulligan's in support of the answer that Marty's work may well have exerted an indirect influence on Grice. §4 ("Commonalities") explored some similarities and differences between Marty's work and Grice's that are differentially visible at different levels of resolution, and suggested that the central insights that are present in Grice's essay are absent from Marty's work. Full assessment of that claim will be dependent on further reflection on the extent to which the purported insights in Grice's essay are genuine. §5 ("Expression") considered an apparent problem that Grice presented for the natural idea that intentional behaviour can manifest or express bits of psychology other than their productive intention. We saw there that one important difference between Marty and Grice centres on their different reactions to this issue. While Grice treated it as a genuine problem for the natural idea of intentional expression, we saw that Marty treated the apparent problem as merely apparent.

My aim has been to suggest that one thing we can learn from Marty—and from historical reflection more generally—is that claims that might otherwise have struck us as uncontestable are contestable, since contested. We may have as much to learn from reflecting on differences amongst thinkers in our subject's history as we do from reflecting on their commonalities.

References

Austin, J. L. 1946. 'Other Minds.' *Proceedings of the Aristotelian Society*, Supplementary Volume 20: 148–187. Reprinted in Austin 1979.

Austin, J. L. 1962. *How to Do Things with Words*. 2nd edn., M. Sbisà and J. O. Urmson eds. Oxford: Oxford University Press, 1975.

Austin, J. L. 1979. *Collected Papers*. 3rd edn., J. O. Urmson and G. J. Warnock eds. Oxford: Oxford University Press.

Cesalli, L. 2013. 'Anton Marty's intentionalist theory of meaning.' In D. Fisette and G. Frèchette eds. *Themes from Brentano*. Amsterdam: Rodolpi.

Gardiner, A. 1922. "The definition of the word and the sentence." *British Journal of Psychology*, XII, 4: 352–361.

Gingerich, O. 1985. 'Did Copernicus owe a debt to Aristarchus?' *Journal for the History of Astronomy* 16, 1: 39–42.

Grice, P. 1957. "Meaning." *The Philosophical Review*, 66: 377–88. Reprinted in Grice 1989. References to reprint.

Grice, P. 1982. 'Meaning Revisited.' In N. V. Smith ed. *Mutual Knowledge*. New York: Academic Press. Reprinted in Grice 1989. References to reprint.

Grice, P. 1989. *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.

Liedtke, F. 1990. 'Meaning and expression: Marty and Grice on intentional semantics.' In K. Mulligan ed. *Mind, Meaning and Metaphysics: The Philosophy and Theory of Language of Anton Marty*. Dordrecht: Kluwer Academic Publishers.

Marty, A. 1908. *Untersuchungen zur Grundlegung der allgemeinen Grammatik und Sprachphilosophie*. Vol. I. Halle: Max Niemeyer.

Marty, A. 1918. 'Über subjektlose Sätze und das Verhältnis der Grammatik zu Logik und Psychologie.' In J. Eisenmeier, A. Kastil eds. *Gesammelte Schriften* Bd. II.1. Halle: Max Niemeyer.

Moran, R. 2005. 'Getting told and being believed.' *Philosophers' Imprint* 5, 5: 1–29.

Ogden, K. and Richards, I. A. 1923. *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. New York: Harcourt, Brace & World.

Rollinger, R. 2014. 'Anton Marty.' *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta ed. URL = <http://plato.stanford.edu/archives/spr2014/entries/marty/>.

Stevenson, C. L. 1944. *Ethics and Language*. New Haven: Yale University Press.