Returning to a Tension within Grice's Original Account of Nonnatural Meaning

KONSTANTY KUZMA[1]

ABSTRACT: It has become a commonplace to regard Grice's project in "Meaning" as plagued by circularity, and almost as prevalent to dismiss such charges as unfounded. Much of the controversy surrounding Grice's presumed circularity revolves around the question whether Grice is committed to a reductionist project of meaning, or whether it is merely meant to elucidate the nature of meaning without pretending to reduce it to something meaningless. Rarely, however, are these views developed as part of a systematic analysis of Grice's original paper, as this paper seeks to do. My paper consists of two parts. In the first part, I try to show how Grice can be defended from John Searle's criticism relating to the famous American soldier example and argue that Searle's suggested amendments run counter to Grice's ambitions. In the second part of my paper, I illustrate – drawing on the first part – why "Meaning" both makes it necessary and seem impossible that the timeless meaning of utterances be fully reducible to individual utterances and thus to individual speakers' intentions. I argue that this seriously challenges the view that Grice is putting forward a theory of intention-based semantics in "Meaning" which would present a viable alternative to later developments of his theory.

KEYWORDS: Grice – intentions – meaning – pragmatics.

✉ Konstanty Kuzma

Department of Philosophy
Ludwig Maximilian University Munich
Geschwister-Scholl-Platz 1 (A 221), 80539 Munich, Germany

e-mail: k.kuzma@campus.lmu.de

0. Introduction

This paper is first and foremost intended to return to and bring out a tension within Grice's seminal "Meaning". While the tension has previously been observed (it is explicitly stated in Strawson 1971 and Burge 1979), it is rarely formulated within the context of a systematic treatment of Grice's original theory of meaning, and to my mind never against the background of its wide implications. There seem to be two principal reasons for this situation. One is that Grice and Schiffer soon developed an alternative way of pursuing a broadly Gricean approach that does not rely on the self-referential intention of "Meaning" and is widely regarded to be the more promising path towards constructing an intention-based semantics (see Grice 1989f; Schiffer 1972). The other is that Grice has often been disassociated from the attempt to fully reduce the semantic to the psychological. As Avramides has argued at length, one can conceive of Grice's project as one of mutual elucidation rather than one-way reduction (cf. Avramides 1989, ch. 1). With this possibility in mind, one can return to Grice's original account from "Meaning" without falling prey to the tension I am about to delve into.

My paper does not pursue either of these approaches. In fact, it is orthogonal to the adequacy of the Schiffer/late Grice approach, and only relevant to the anti-reductionist insofar as it (implicitly) disassociates her from the project of an intention-based semantics. This is because the paper puts pressure on the third alternative the above landscape leaves open, which is to return to the original account of "Meaning" with a reductionist project in mind.[2] I will argue that this is both the most natural way to read that original paper (Section 1), and the only hope of reconstructing "Meaning" as an account of intention-based semantics (Section 2). Because of this limited perspective on Grice's larger project, I will mainly draw on John Searle, who has argued that the self-referential intention from "Meaning" is key to a proper account of meaning (cf. Searle 2007, 14).[3] My aim will be to show that Grice's original approach is inconsistent as an attempt

---

[2]   Thus, I pursue what Avramides refers to as a "strong, reductive interpretation" (Avramides 1989, 13).

[3]   This is so even though Searle also thinks Grice confuses the explanatory role self-referentiality should play in such an account.

to construct an intention-based semantics, so that pace Searle, we should not regard the self-reflexive intention as key to the Gricean project.[4]

The tension my paper is concerned with is the following. As I will argue, "Meaning" both makes it necessary and seem impossible that the timeless meaning of utterances be fully reducible to individual utterances and thus of individual speaker's intentions.[5] This is because for Grice to provide a theory of meaning – specifically, an intention-based semantics –, the explicatory dependency between timeless meaning and speaker's meaning must be one-way (this is the requirement meant to be brought out by the mildly dramatic talk of "full" reducibility as opposed to reducibility simpliciter above).[6] That is, for Grice's project to succeed as an instance of intention-based semantics, timeless meaning must be analyzable in terms of speaker's meaning without semantic remainder.[7] The fact that "That book is green" means what it means, for instance, must solely be a function of a community of speakers intending it to mean what it means. Once this requirement is brought into view, another difficulty arises, which is that there seems to be no way of meaning anything complex by one's utterances independent of the existence of timeless meaning. For example, there is no hope of meaning that the book over there is green without there being a set of conventions which fix the

---

[4]   The self-reflexive intention of "Meaning" has been confronted with concerns about its presumed circularity and implausibility, prompting the development of alternative approaches to the intended effect of an utterance. Cf. Neale (1992, 548); Recanati (1986); Sperber & Wilson (1986). I will argue that even if we grant Grice that there is nothing circular or implausible about the self-reflexive intention, one cannot both hold on to Grice's original account from "Meaning" and pursue the project of an intention-based semantics. It is in this sense that the paper is meant to discredit Grice's original account as a viable intention-based alternative to later versions of the theory.

[5]   Strawson acknowledges this tension but thinks that you need not posit full reducibility from a Gricean perspective (cf. Strawson 1971, 174). I discuss Strawson's solution in Section 2.

[6]   Drawing on Grice's characterization of both terms, I use "speaker's meaning" to denote the meaning intended by the speaker in uttering an utterance, whereas "timeless meaning" denotes the conventional meaning of an utterance.

[7]   I borrow this way of framing the requirements of Grice's theory from Grandy & Warner (2017). This sets my reading of Grice's original paper apart from interpretations that take Grice to merely aspire a reductionist project in the sense of conceptual elaboration (see e.g. Neale 1992).

meaning of a set of signs roughly synonymous to "That book is green". And that would of course mean that whatever the account presented in "Meaning" amounted to, it would not serve the reductionist project that Grice is aiming at. For that, again, would require timeless meaning to be fully reducible to speaker's meaning without semantic remainder.[8]

I will approach the said tension through a discussion of two lines of criticisms that Searle has raised vis-à-vis Grice's conception of meaning. Though I share Searle's verdict that Grice fails to provide a theory of meaning, I share it for different reasons, meaning that the discussion of Searle will lead up to my criticism of Grice in a roundabout fashion. I will first rehearse Searle's criticism of Grice, then try and show how Grice can be defended against it, and finally argue that Searle's objections and his American soldier example (to which I will get shortly) can nevertheless help us see what is fundamentally problematic with the conception that Grice offers in "Meaning".

1. Searle's criticism

I will begin my discussion of John Searle's criticism of Grice with the arguments put forward in *Speech Acts*. Though Searle later revised the position argued for in *Speech Acts*, it will be helpful to briefly return to it. The object of Searle's inquiry in *Speech Acts*, then, is the central definition of nonnatural meaning argued for in "Meaning", which Searle cites in the following fashion:

> To say that a speaker *S* meant something by *X* is to say that *S* intended the utterance of *X* to produce some effect in a hearer *H* by means of the recognition of this intention. (Searle 1969, 43)

---

[8] I stress Grandy and Warner's "without semantic remainder" proviso (see Grandy & Warner 2017) because there is an obvious sense in which Grice reduces all meaning to intentions. After all, speaker's meaning is constituted by intentions, while timeless meaning is nothing but regularities among those very intentions. The issue I will discuss towards the end of section 2 is that while intentions determine the individual meaning of utterances, they can only do so against the context of an already existing set of conventions.

It will be useful to have a shorthand for this definition, so let us call it Meaning[NN]. The purpose of Meaning[NN]'s self-reflexivity ("by means of the recognition of this intention") is to separate cases of nonnaturally meaning something from cases in which one intentionally produces an effect in someone without one's intention playing any part in the production of that effect. The latter case, Grice argues, would not be a case of nonnaturally meaning anything. An example for this case cited in "Meaning" is the scene of Herod presenting Salome with the head of John the Baptist (Grice 1989b, 218). While Herod intended to make Salome believe that John the Baptist has died by producing the latter's head on a platter, it is not the case (or so argues Grice) that Herod *meant* anything by showing her the head of Salome. Herod's intention to make Salome believe that John the Baptist is dead does not play a role in producing the effect of her believing that John the Baptist is dead. If, on the other hand, Herod had (to the unquestionable detriment of art history) relied on less dramatic means of getting the message across and simply *said*, "I had John the Baptist killed", his intention of getting her to know that John the Baptist is dead would have indeed played a part in producing that effect. Therefore, the latter case would have been a case of nonnatural meaning.[9]

Searle takes issue with this account for two reasons. One is that Grice does not account for the way that meaning "can be a matter of rules or conventions" (Searle 1969, 43). In other words, Searle claims that "Meaning" does not acknowledge the way in which meaning something by one's utterance is connected to what that utterance usually means. The other is that Grice is supposed by Searle to be wrong about the intended effect of utterances. While "Meaning" states that nonnaturally meaning something by one's utterance (in the case of indicative sentences) is an instance of intending to "induce by x a belief in an audience" (Grice 1989b, 219), Searle thinks that it is merely an instance of producing understanding on the hearer's part. Since the latter objection is developed in Searle's recent paper "Grice on Meaning: 50 Years Later", the discussion of which I will

---

[9]   It has been debated whether Grice's intuitions are correct concerning the contrast between natural and nonnatural meaning in the Herod example. In particular, it is controversial whether the self-reflexive clause (which posits that the intention to produce an effect must itself be intended to function as a reason for producing that effect) is needed. Cf. Neale (1992, 548); Recanati (1986); Sperber & Wilson (1986).

take up shortly, I will – for the time being – concentrate on the first objection, i.e. that Grice's account fails to account for the connection between speaker's meaning and timeless meaning, which brings us to Searle's famous American soldier example.[10]

The example goes as follows. We are supposed to imagine that an American soldier who has been captured by Italian troops is trying to make his captors believe that he is a German officer. Knowing virtually no Italian or German, he puts on a show to tell them that he's a German officer by reciting the only German line that he knows: "*Kennst du das Land, wo die Zitronen blühen?*" Searle maintains that the soldier's utterance does not mean either "I am a German officer" or that utterance's German-language equivalent "*Ich bin ein deutscher Offizier*". But the Gricean analysis, he thinks, not only implies that this is what it means, but that furthermore it follows that "any sentence can be uttered with any meaning whatever, given that the circumstances make possible the appropriate intentions" (Searle 1969, 45). To prevent meaning from being fixed arbitrarily, Searle suggests incorporating the conventional meaning of utterances into Grice's account of meaning. Thus, Searle arrives at the following, amended version of Grice's account of meaning:

> In our analysis of illocutionary acts, we must capture both the intentional and the conventional aspects and especially the relationship between them. In the performance of an illocutionary act in the literal utterance of a sentence, the speaker intends to produce a certain effect by means of getting the hearer to recognize his intention to produce that effect; and furthermore, if he is using words literally, he intends this recognition to be achieved in virtue of the fact that the rules for using the expressions he utters associate the expression with the production of that effect. It is this *combination* of elements which we shall need to express in our analysis of the illocutionary act. (Searle 1969, 45)

---

[10] The example is presented in Searle (1969, 44f). Notable (and for the most part dismissive) discussions of the example are to be found in Grice (1989f); Armstrong (1971, 440-441); Bennett (1973, 164-165); Martinich (1984, 122-125); Schiffer (1972, ch. 2); Yu (1979, sct. 3).

This establishes the connection between speaker's meaning and timeless meaning that Searle's above-mentioned criticism of Meaning^{NN} called for. Meaning something by one's utterance is not a completely arbitrary bestowal of meaning on an utterance that can by that act be made to mean anything. Rather, Searle thinks, "what we can mean is at least sometimes a function of what we are saying" (Searle 1969, 45). Citing Wittgenstein, Searle reminds us that you cannot say "it's cold here" and *mean* the opposite (Searle 1969, 45).

There are several ways of responding to Searle's criticism, some of which Grice himself pointed to in "Utterer's Meaning and Intentions". Before I discuss some of those responses, however, it is worth pausing for one moment to deliberate Searle's counter-suggestion which he thinks provides a way of both avoiding counter-examples of the sort exemplified by the American soldier and establishing a connection between speaker's meaning and timeless meaning. Even if we set aside the problem posed by counter-examples for a moment, it is quite clear that Searle's suggestion for amending Grice's account of meaning will not do as far as Grice's project is concerned. This is because the connection that Searle establishes between an utterance and its conventional meaning makes it impossible to arrive at a reductive account of meaning. Searle suggests that literal utterances be thought of as resulting from a combination of the utterer's intention and his utterance's conventional meaning, so that the analysis of meaning includes the very thing that is supposed to be explained. In other words, we have arrived at an explanation of meaning which itself includes a reference to meaning in the form of "rules for using the expressions he [the speaker] uses". Initially, it is not entirely clear *what* kind of nonnatural meaning Searle is attempting to give an account of – whether it is of an utterance's timeless meaning or of speaker's meaning. But in the former case, the account would be blatantly circular, as Searle would be analyzing an utterance's timeless meaning in terms of the utterer's intentions and the utterance's timeless meaning. Even if we are more charitable towards Searle and allow that he is attempting to provide a definition for speaker's meaning, whereas the "rules for using the expressions" are clearly a reference to timeless meaning, the problem remains standing that in trying to account for meaning, he is making recourse to something that is already meaningful (namely the rules for expression use). While not strictly speaking circular, Searle would still not be providing a proper account of what makes

utterances meaningful. If we have conventional meaning to fall back on in uttering literal utterances, we do not need intentions to make them meaningful – they already are.[11]

It will be useful to keep this in mind as Searle's misconception of Grice's aim in "Meaning" shapes his entire discussion of it. I now turn back to the American soldier example to show how Grice can deal with it. To recap, the example was supposed to pose a problem for Grice because his analysis would suggest that "*Kennst du das Land, wo die Zitronen blühen?*" could be brought to mean "I am a German officer", whereas of course it means something completely different. Now as Grice rightly points out, it is quite difficult to imagine how the American soldier could want to get his captors to think that the *words* he is uttering mean "I am a German officer" (Grice 1989f, 101-102).[12] It would be much more natural to describe the example analogously to the Herod case, so that the Italian troops merely infer, from the observed circumstances of the soldier's utterance (i.e. independently of his intentions) – his German-sounding words and the fact that he has the demeanor of a representative of the military – that he is a German officer.[13] And if this were indeed the proper way to describe the scenario, then (again analogously to the Herod example) we would not be dealing with a case of nonnatural meaning. For even if the American soldier had intended his captors to go through the said inferential steps, his intention could not have been supposed to play a role in their arriving at the belief that he is a German officer.

The argument could have ended here. As Grice himself observes, this is the most intuitive way of describing Searle's example, and Grice's response seems both ample and satisfactory. But charitable as he is, Grice

---

[11]   It is worth noting that intentions appear to do no work on the latter reading of Searle's account of the meaning of literal utterances. For what does it matter what *I* mean by a literal utterance if its meaning is already fixed by the conventional meaning? Martinich argues that because of such constraints, Searle effectively ties utterances to their conventional meaning (Martinich 1984, 124).

[12]   The same issue is raised in Armstrong (1971, 440); Bennett (1973, 164). Schiffer goes even further in questioning whether the American soldier meant anything at all by his utterance. See Schiffer (1972, 27).

[13]   This is the first of two possible reinterpretations of Searle's example offered in Schiffer (1972, 28). Also see Grice (1989f, 101).

allows that Searle have his way. He assumes with Searle that the American soldier in fact wants the Italian troops to come to believe that he is a German officer "via a belief that the words which he uttered were the German for 'I am a German officer'" (Grice 1989f, 101). And if this is something we can imagine, Grice continues, then we should say that the American soldier meant by "*Kennst du das Land, wo die Zitronen blühen?*" that he is a German officer. Does this sound counter-intuitive? Hardly so, for Grice explicitly denies the implication that this is what the German officer means by the *words* "*Kennst du das Land, wo die Zitronen blühen?*" (Grice 1989f, 102). The relevant analysis presented in "Meaning" is merely intended to bring out what a Speaker S means in *uttering* a sentence X. There is no good inference to the commitment on Grice's part that that is what the sentence normally means. And of course, we can mean something in making an utterance that departs from its conventional meaning. One need not refer to Grice's theory of conversational implicature to acknowledge this point. Even someone who objected to that theory's logical ramifications must surely acknowledge that departure from the normal meaning of one's utterances is something we do on a daily basis.

Imagine the following scenario: a group of friends meet in a bar to have a drink. When the waiter comes by to take everyone's order, one of the friends misspeaks while ordering a beer, saying "bear" instead of "beer". The group breaks out in laughter, and the waiter picks up the slip of the tongue, reacting with a dry joke which produces further laughter among the group. The next time that the waiter comes by to take orders, the friends order "bears" rather than "beers". It seems quite natural to describe the friends as intending to order "beers" when asking to get "bears" from that moment on, and to expect the waiter to understand their cue. Still, there is no good inference to saying that this is what the word "bear" means. In fact, it is precisely due to it not being the word's conventional meaning that it provokes laughter among the group.

The mistake on Searle's part is to assume that Meaning^{NN} is supposed to do all the work for Grice's theory of meaning. When Grice is saying that in uttering a sentence X, S intends to produce an effect in hearer H in virtue of H's recognition of that intention, he is giving an account of speaker's meaning. In other words, he is trying to give an account of what happens when someone means something by an utterance. But this is not to say that a speaker's meaning something by an utterance fully accounts for the

utterance's meaning. In particular, it does not account for the utterance's timeless meaning. Speaker's meaning only accounts for what a speaker means by uttering an utterance.

In certain cases, this may be all that matters. In the American soldier example (assuming that it is an instance of nonnatural meaning), it is quite irrelevant that the sentence uttered by the soldier really means "Knowest thou the land where the lemon trees bloom?", as its conventional meaning stands in no relation to the meaning the speaker intended to convey in uttering that sentence in German. But in most cases of verbal communication, the command of the spoken language will be essential to deciphering the intended message. As Grice illustrates in "Logic and Conversation", this is even true of cases of nonliteral speech, as the ability to understand its meaning requires an understanding of the literal meaning of an utterance just as much as instances of literal speech do (Cf. Grice, 1989a, 30f). It is only through an understanding of an utterance's literal meaning that one arrives at the intended meaning of a non-literal utterance, for one must first pass through it and its inadequacy when interpreted literally (together with an application of the Cooperative Principle and possibly the Maxims of Conversation) to be able, according to Grice, to arrive at a proper understanding of how to take the message instead.[14]

This is also why, pace Searle, one cannot arbitrarily fix the meaning of utterances. Even when an utterance is meant non-literally, the literal meaning of the utterance plays a part in arriving at its meaning.[15] Grice is very explicit on the so-called M-intentions (that is meaning intentions) being bound by what can be transferred in an act of communication (this is part of the reason why he repeatedly refers to conversations as a rational

---

[14]   As has been pointed out to me, the literal adequacy of certain metaphorical utterances (e.g. "No man is an island") puts pressure on the view that one should regard the inadequacy of an utterance when interpreted literally as a necessary point of departure for interpreting non-literal utterances. Arguably, "inadequacy" may thus be too strong a word, though the question of the exact formulation of the process of getting from literal to non-literal meaning is not entirely relevant to the purpose of this paper. The important point is that communication is rationally constrained (in part by timeless meaning).

[15]   This is the principal reason why Neale objects to the claim that Grice neglects the role that timeless meaning plays in working out communicative intentions. Cf. Neale (1992, sct. 6).

endeavor) (cf. Grice, 1989a, 31). It is not the case that one can utter, in the middle of an unrelated conversation, "blob" with the intention of producing in one's hearer the belief that "Lewis Carroll is the pen name of Charles Lutwidge Dodgson". Rather, one can mean by one's utterances what they normally mean as well as what can be conveyed by them non-literally by way of drawing upon the principles of conversation laid out in "Logic and Conversation". And this bars one from being able to mean anything by anything whatsoever.[16]

There is of course a trivial sense in which it is indeed the case that one can mean anything by anything whatsoever, namely that we can imagine circumstances under which any sentence could be used to mean just about anything. Surely, we do not want to deny that under specific circumstances, say during a game, we could stipulate or otherwise imply that utterances mean something else than what they normally mean and be almost unlimited in our freedom to do so. But this does not imply that you could mean anything by uttering any sentence solely in virtue of your intending it to mean something, as in the Lewis Carrol example above, nor does it imply that this is what the words the utterance consists of mean. It is true that Searle thinks it is a problem that you can in principle mean anything by anything even though he acknowledges that this is only true given "that the circumstances make possible the appropriate intentions" (Searle 1969, 45). But this is only because Searle a) does not acknowledge the way in which circumstances put a rational constraint on what you can mean by an utterance, and because b) he does not realize that Grice is speaking about an utterance's speaker's meaning, and not its timeless meaning (cf. Armstrong 1971, 440-441; Bennett 1973, 164-165). Once we appreciate the latter two reservations, the charge becomes harmless. Grice is not committed to the absurd view that you can, under any given circumstances, mean "Lewis Carroll is the pen name of Charles Lutwidge Dodgson" by uttering anything whatsoever. His view is rather that the obtaining of appropriate circumstances allows you to utter that sentence while meaning something other than what is normally meant by its utterance. Furthermore, by that act, the sentence will not suddenly change its timeless meaning, but will

---

[16] The extent of this rational constraint is dangerously downplayed in Martinich (1984, 122-125). See Neale (1992, especially scts. 5 & 6), for an exposition of Grice's need for and deliverance of rational constraint on what an utterer can mean.

instead help instantiate a different utterance's meaning than that which is normally associated with its utterance.

This addresses the first line of critique raised in *Speech Acts*. It is not the case that Grice makes no connection between utterances and their time-less meaning (or conventional meaning), as one in general needs to know the literal meaning of utterances to be able to decipher even their non-literal meaning (as I pointed out above, the American soldier example – when interpreted as a case of nonnatural meaning – is a notable exception). This is why Grice repeatedly stresses that one is generally assumed to be intend-ing to convey the literal meaning of one's utterances, which assumption is only dropped if it cannot be reconciled with the speaker's observing the Cooperative Principle (cf. Grice 1989b, 222; Grice 1989a, 30f). But even such deviance from the literal meaning of one's utterances and the arriving at their non-literal meaning is rule-governed, so that Searle is wrong in claiming that his American soldier example shows that one can, on a Gri-cean picture, mean anything by uttering any sentence whatsoever (lest we mean by that that we can imagine appropriate circumstances under which any given utterance could be brought to mean anything, which implication, for the reasons mentioned above, would not be problematic).

What about the second line of critique raised by Searle in *Speech Acts*, i.e. the charge that Grice is wrong about the intended effect of utterances? As I already mentioned, Searle altered and amended his critique from *Speech Acts* in his recent paper "Grice on Meaning: 50 Years Later", so that it will be useful to look at both texts to get an idea of Searle's position. To reiterate, in *Speech Acts* Searle objected to Grice's contention that nonnaturally meaning something by an utterance is an instance of trying to induce a belief in an audience. As Searle points out, we can utter a sentence and nonnaturally mean something by it without having any intention of inducing a belief in our audience (Searle 1969, 46-48). Even if we restrict our attention to indicative sentences, an analysis of which Grice was chiefly attempting to give in "Meaning", this charge seems justified. Take the fol-lowing promenade example: if I take a stroll with a friend and remark on the beauty of the surroundings by saying "What a remarkable landscape this is!", it seems wrong to suggest that I am thereby attempting to get my audience to *believe* that the landscape is remarkable. Nor does it appear right to say, as Grice later claimed when responding to criticisms of this sort (Grice 1989e, 123), that I am trying to inform my friend of *my own*

belief about the remarkableness of the landscape.[17] Indeed, it seems wrong to describe my utterance as being primarily concerned with beliefs at all.[18]

Searle's suggestion in *Speech Acts* was that the proper way to describe someone's nonnaturally meaning something by an utterance is to say that the speaker is thereby intending to produce understanding on the hearer's part. Hence, in the promenade example I am merely getting my friend to understand what it is I am trying to say. In Searle's analysis, there is no reference to beliefs anymore, but only to getting my hearer to know what it is I am trying to communicate. But Searle later admitted that it was possible to nonnaturally mean something by one's utterance without even intending to produce understanding in one's audience (Searle 2007, 13f). A standard example is soliloquy, which also does not seem to be an example of producing understanding in one's hearer even though it is clear that one nevertheless nonnaturally means something by one's utterances.[19] The lesson Searle draws from this in "Grice on Meaning: 50 Years Later" is that Grice confused his account for an account of meaning, when in fact he was giving a (flawed) account of communication, and that Searle's analysis in *Speech Acts* followed Grice in this. Importantly, Searle still contends in "Grice on Meaning" that his *Speech Acts* account is superior to that of Grice because it can deal with all cases of nonnatural meaning save for

---

[17]   Of course, Grice's attempt to save his analysis by suggesting that it is one's own beliefs that one is attempting to convey when nonnaturally meaning something by an utterance can best be challenged by finding examples in which one *is* in fact attempting to get the audience to believe something, as Grice's earlier analysis suggested was always the case. If, for instance, I am having a lengthy argument about youth unemployment in Europe with a friend, and after half an hour I verbally produce a statistic which I think will be devastating for my friend's position, it is clear that in producing the statistic I am not getting him to think what my position is.

[18]   There is extensive literature on the problem of conceptualizing an utterance's intended effect. See Lycan (2008, 89-91); Neale (1992, sct. 5); Schiffer (1972, ch. 3); Strawson (1971, 172-173).

[19]   See Ziff (1967) and Vlach (1981, 384-386), for useful expositions why audienceless cases constitute a problem for Grice (and, by extension, for Searle as well). Grice invokes the audience counterfactually to deal with the problem (so that one should understand the utterer as intending that were there an audience, the intended effect would be brought about), while Schiffer argues that at least in certain cases of soliloquy the speaker is his or her own audience. Cf. Grice (1989f, sct. 5); Schiffer (1972, 80).

those in which the speech act performed is defective (for lack of an intention to produce understanding), whereas there are perfectly nondefective speech acts (such as the promenade example) which Grice's analysis cannot deal with. The fact remains, however, that the accounts given in "Meaning" and *Speech Acts* are unable to deal with cases in which someone nonnaturally means something by one's utterance without intending to produce any effect in one's audience, and that according to Searle they provide accounts of communication rather than meaning as a result (Searle 2007, 14).

Given Searle's above-mentioned confusion about the aim of "Meaning", it is no surprise that his revised account of meaning in "Grice on Meaning" again fails to provide the kind of theory of meaning that Grice was looking for. Searle's new suggestion is to think of the literal meaning of indicative sentences in truth-functional terms:

> The meaning intention consists in the *intentional imposition of conditions of satisfaction* (in the sense of requirement) on conditions of satisfaction (in the sense of things required). The initial condition of satisfaction is simply that I produce the utterance, but the distinction between the utterance without meaning it, and the meaningful utterance where the meaningfulness is intended, is that the utterance itself, the condition of satisfaction of my intention to produce that utterance, has further conditions of satisfaction. In this case [i.e. a literal, indicative sentence] it has truth conditions. […] Analogous remarks can be made about directives and other forms of speech acts. Thus if I utter the French sentence "Fermez la porte" without meaning it, but just, for example, as practicing French pronunciation, the condition of satisfaction of my intention in action is simply that the intention in action should produce that utterance. But if I not only utter it but mean it, that is, mean it as a directive, then the conditions of satisfaction include that the hearer close the door. (Searle 2007, 15f)

We can leave aside any worries about whether one can really distinguish as easily between cases of saying and meaning something and saying something without meaning it. The important thing is that the impossibility of producing a reductive account of meaning reemerges. For Searle again explains meaning in terms of an act which is already meaningful, namely the intentional imposition of truth functions on an utterance. The articulation

and imposition of truth conditions on a series of sounds requires that those sounds be adequate vehicles for the transfer of those truth conditions. Otherwise the problem of being able to mean literally anything by any utterance whatsoever would emerge, as one could indeed be able to utter "blob" and mean "Lewis Carroll is the pen name of Charles Lutwidge Dodgson" just in virtue of one's wanting it to mean just that. In truth, we find that we are dependent on the timeless meaning of utterances for them to have truth conditions. "It is raining" is true if it is raining because that is what it means, and not just in virtue of my wanting it to be true under those conditions of satisfaction. In other words, Searle again ends up explaining meaning based on something that already presupposes meaning, namely the intentional imposition of truth functions on an utterance.

To be fair, in "Grice on Meaning" Searle does not pretend to be fixing Grice's account of meaning, but instead declares that his account is "Gricean in spirit" (Searle 2007, 17). But even his contention that Grice mistakes his account for a theory of meaning is to be handled with care. As I already hinted at, it is not the case that Meaning[NN] is supposed to do all the work on Grice's account of meaning. That is why, granted that it may be better to describe the intention behind nonnaturally meaning something by an utterance as wanting to produce understanding rather than belief or belief communication, it is no real threat to Grice's project to state that Meaning[NN] provides an account of "communication" as opposed to "meaning", so long as one means by this that Grice is giving an account of what it means for someone to nonnaturally mean something by an utterance, and not, as Searle wrongly suggests, an account of what utterances normally mean (more on this below).

2. Grice's account and a fundamental weakness

We said that Grice's account of meaning could be summarized in the following fashion: to say that a speaker S meant something by X is to say that S intended the utterance of X to produce – to incorporate Searle's suggestion – understanding in a hearer H by means of the recognition of this intention. This, again, is meant to account for what a speaker means in uttering an utterance. In other words, Grice is presenting an account of speaker's meaning.

As we saw, speaker's meaning does not necessarily correspond with what the uttered words normally mean. Indeed, you can nonnaturally mean something by uttering a series of sounds that has no conventional meaning in any language whatsoever (even though this requires that appropriate circumstances obtain). Normally however, you do need a grasp of the literal and conventional meaning of a sentence to unpack the utterance – whether it is literal, or not. If Searle were right and Grice were attempting to account for meaning solely in virtue of Meaning[NN], then Grice too would not be offering a reductive account of meaning. For even if we granted Grice that one cannot mean anything whatsoever by any odd utterance but must instead arrive at an understanding of speaker's meaning by way of the utterance's literal meaning, his account would fail as a theory of meaning for the same reasons as Searle's counter-suggestions did if he simply presupposed the existence of literal meaning.

But Grice has a separate story to tell about timeless meaning, which he thinks is indeed reductive. This is that the timeless meaning of an utterance is a function of what speakers in a linguistic community mean by that utterance, a view Brandom has called "regularism" (Brandom 1994, 26-30). According to regularism, rules guiding the use of an expression – on a rule-based account, its meaning – are nothing other than a description of the regularities pertaining to the use of those expressions within a linguistic community. In other words, to talk about norms "is just to talk about regularities" (Brandom 1994, 27), a view which Kripke famously attributes to Wittgenstein (cf. Kripke 1982, sct. 3). This yields a complex picture of Grice's project of an intention-based semantics. On the level of individual utterances, an utterance means what a speaker intends it to mean. But it is regularities among just this kind of M-intentions which yield the timeless meaning of utterances. And it is of course the latter which figure as a constraint on what speakers can intend by individual utterances.

Does this important aspect of Grice's theory save the project of an intention-based semantics as presented in "Meaning"? It would if it allowed us to reduce timeless meaning to speaker's meaning without semantic remainder. But as the discussion of Searle's American soldier example brought out, Grice's account of speaker's meaning cannot be conceptualized independently of timeless meaning. Recall that for a speaker to nonnaturally mean something by his utterance, and for a hearer to understand it, both usually rely on the literal meaning of the utterance (a notable

exception are utterances which are not based on or even related to actual words). Thus, if I make the utterance "It is raining", my hearer will normally have to know what the utterance usually means to be able to understand what I mean by uttering it.[20] This is evidently so in the case of literal utterances, because the timeless meaning of "It is raining" would then be precisely what I am trying to communicate. But it is also true if my utterance were meant non-literally, in which case the hearer would have to draw upon the Cooperative Principle (and possibly the Maxims of Conversation) to infer, according to the principles laid down in "Logic and Conversation" what I mean in uttering the sentence "It is raining". As I laid out above, this is the reason why speakers are generally assumed to be intending to mean their utterances literally. Now this reliance on the literal meaning of utterances brings up the same problem that was earlier put forward against Searle. For a speaker to be able to mean something by his utterance, he relies on the timeless meaning of it independently of the question whether it generally communicates what he is intending to convey. It does not matter that the timeless meaning of utterances derives from members of a linguistic community intending to mean something by it. For those speakers, in having meant something by their utterances, themselves relied on timeless meaning in uttering the utterances which contributed to the utterance's meaning what it means today. So that Grice, like Searle, is unable to explain meaning without presupposing something that is itself meaningful – namely timeless meaning. To be able to mean anything, one must already be able to draw on the proper vehicles for communicating that meaning. Intentions cannot do that work on their own.

How does this relate to Searle's criticism and his American soldier example? I think that Searle's objections and his example in particular bring out very clearly why we cannot have meaning reduce to intentions without relying on some form of conventional meaning. Of course, according to my analysis, this follows from the objections and the American soldier example in a roundabout way, since I agreed with Grice that the most natural way to describe the example is as a case of natural meaning. It is quite

---

[20]   As Neale rightly points out, this does not undermine the idea that what an utterer means is determined by his communicative intentions. Cf. Neale (1992, 553). It does, however, put a rational constraint on what the (semantic) preconditions of communication are.

difficult to imagine how the American soldier should have brought his captors to believe that the *words* "*Kennst du das Land, wo die Zitronen blühen?*" mean "I am a German officer". But it is precisely this difficulty that forces us to deliberate *why* the American soldier cannot just get the soldiers to think that that is what those words mean. And the answer is that one cannot mean anything by anything whatsoever when hearer and speaker do not have a shared grasp of timeless meaning to fall back on, as happens when one is trying to speak to someone without having a shared language to mediate. Ironically, this is something made excessively clear by Grice's myth about the presumed origins of language, through which Grice inadvertently reveals how powerless (if conceivable at all) intentions are when there are no conventions to fall back on (Grice 1989c, 290-297).

The idea of autonomous intentions further recedes when one recalls how Grice himself describes the process of communication. We said that a hearer must in general grasp the literal meaning of an utterance to be able to decipher its speaker's meaning, and that the decision whether a given utterance is literal, as well as its re-interpretation in cases of non-literal speech, follows rational principles which were laid out in "Logic and Conversation". In addition to these principles and the utterance's general usage, a hearer intending to interpret an utterance can, according to Grice, (sometimes) rely on explicitly formulated linguistic (or quasilinguistic) intentions, the context of the utterance (linguistic or otherwise) or, in difficult cases, a deduction to determine the speaker's meaning (cf. Grice 1989b, 222-223; Neale 1992, sct. 6). Now it is striking that there is recourse to intentions only in the case of the explicitly formulated intentions, which are introduced with the caveat that they are *not* conclusive. According to Grice, "a speaker who has declared an intention to use a familiar expression in an unfamiliar way may slip into the familiar use" (Grice 1989b, 222). In other words, even when someone announces explicitly how to take his words, his deeds determine the outcome. Not only are intentions dependent on timeless meaning to even be articulated (let alone communicated). Even after having been formed *and* explicitly verbalized, timeless meaning still serves as an interpretive device that helps determine whether any given utterance was really intended in the way that the utterer has claimed it is. So that timeless meaning is both a presupposition for the functioning of intentional communication, and a retroactive corrective.

None of this is to say that intentions play no role in the process of communication, nor even that they play no role in constituting an utterance's meaning. Firstly, to ask what is meant by an utterance is to ask what a speaker *meant* by uttering a sentence, even when that sentence is intended by her to be taken literally. In other words, on a Gricean picture it is the speaker's intentions one is after when working out the meaning of an utterance U even in those cases where its vehicle is timeless meaning. Secondly, in working out the meaning of an utterance U that is not intended literally, one will only be able to understand the speaker by trying to figure out how she *meant* her utterance (which the hearer can do by taking the speaker to observe the Cooperative Principle and by applying the interpretation procedure sketched above). Thirdly, on a Gricean picture intentions play a role in the genesis of timeless meaning, insofar as their successful transfer once made it possible to develop a language and through it a means of communicating (and perhaps even forming) complex intentions.

As Neale and Strawson have pointed out, there is thus no circularity or inconsistency in Grice's original project (cf. Neale 1992, scts. 5 & 6; Strawson 1971, 174-175). Even if we concede that, as I have argued, intentions do not get off the ground without some conventions to fall back on, you can explain how intentions fit into the larger picture of the Gricean project without opening the project to definite objections.[21] Still I believe it is often overlooked that the consistency of Grice's project in "Meaning" comes at a price. If we follow Neale and Strawson in having (complex) communication rely on conventions, we distance ourselves from a view according to which intentions are what makes utterances meaningful in the first place. Whether we communicate via literal or nonliteral speech, timeless meaning is needed to work out (and perhaps even form) speaker's intentions. Thus, timeless meaning is a function of a community intending it to mean what it means only in the sense that they determine that individual utterance's meaning, but the fact that they can thus intend it to mean what it means is itself already a function of being able to intend utterances to mean something. If this is the right way to describe Grice's project in "Meaning", it is not concerned with explaining how meaning comes about, but with explaining how *given* our ability to mean things through utterances, individual

---

[21]   See Avramides for a discussion of the advantages of a "weak, nonreductive interpretation of Grice's analysis" (Avramides 1989, 19). Avramides (1989, ch. 1).

utterances come to mean what they mean. In this sense, intentions are not constitutive of meaning because you cannot make sense of communication intentions independent of pre-existing meaning. The project of reducing timeless meaning to speaker's meaning without semantic remainder – which I have treated as the aim and defining feature of any intention-based semantics – cannot be pursued with the tools of "Meaning" alone.[22]

Strawson explicitly addresses this objection in "Meaning and Truth". He agrees that it would be absurd to credit ourselves with "extremely complicated communication-intentions (or at least desires)" independently of "linguistic means of fulfilling those desires" (Strawson 1971, 174). And he does seem to think that a project of the Gricean sort would falter should there be nothing more to say in its favor. That said, Strawson also believes that the project only requires that you can explain conventions of communication "in terms of the notion of pre-conventional communication at a rather basic level" (Strawson 1971, 174). And this is something he deems possible along the lines of Grice's already mentioned genetic account:

> Suppose an utterer achieves a pre-conventional communication success with a given audience by means of an utterance, say *x*. He has a complex intention, *vis-à-vis* the audience of the sort which counts as a communication-intention and succeeds in fulfilling that intention by uttering *x*. Let us suppose that the primary intention was such that the utterer *meant*

---

[22] Thus Neale's establishment of the non-circularity of Grice's project is orthogonal to my concerns. Neale seems to want to establish – pace objections that claim the contrary – "that typically the hearer must establish what *U* has said (or made as if to say) in order to establish what *U* meant; and it is by taking into account the nature and purpose of rational discourse that the hearer is able to progress (via, e.g., conversational implicature) from what *U* has said (or made as if to say) to what *U* meant" (Neale 1992, 552). It is precisely because I agree with Neale in this (see section 1 of this paper) that I see a problem for Grice. Neale does not acknowledge this issue because he appears to be concerned a) with a broadly Gricean approach (rather than "Meaning" taken in isolation), and because he seems to hold that b) Grice can get by without a strong, reductive approach. Since my paper is concerned with showing that (pace Searle) Grice's earlier approach is *not* a viable alternative to his later approaches, I treat "Meaning" in isolation, which no longer leaves open the possibility of holding on to a weak, reductive reading (since the project of "Meaning" builds on meaning fully reducing to intentional states).

that *p* by uttering *x*; and, since, by hypothesis, he achieved a communi-
cation-success, he was so *understood* by his audience. Now if the same
communication-problem presents itself later to the same utterer in rela-
tion to the same audience, the fact, known to both of them, that the ut-
terer meant that *p* by uttering *x* before, gives the utterer a reason for
uttering *x* again and the audience a reason for interpreting the utterance
in the same way as before. (The reason which each has is the knowledge
that the other has the knowledge which he has.) So it is easy to see how
the utterance of *x* could become established as between this utterer and
this audience as a means of meaning that *p*. Because it has worked, it
becomes established; and then it works *because* it is established. And it
is easy to see how this story could be told so as to involve not just a
group of two, but a wider group. So we can have a movement from an
utterer pre-conventionally meaning that *p* by an utterance of *x* to the
utterance-type *x* conventionally meaning that *p* within a group and
thence back to utterer-members of the group meaning that *p* by a token
of the type, but now *in accordance with the conventions*. (Strawson
1971, 174-175)

This is not a *prima facie* implausible account. In fact, one could go so far
as to claim that science gives us evidence of convention-fixing of the above
sort, say among primates. The issue with this solution is rather that its plau-
sibility is seriously strained when one fills in the specific details of Grice's
account, which are of no concern to Strawson in "Meaning and Truth".
Remember that to say that a speaker S meant something by X, according
to Grice's original account that was rehearsed above, is to say that S in-
tended the utterance of X to produce an effect in a hearer H (whether it is
belief or understanding) by means of the recognition of this intention. For
Strawson's argument to work as a defense of Grice's account, Strawson
would be committed to claiming that the utterer in the above example is
not only trying to make his hearers believe something by producing a cue
(as in Grice's Herod example), but that he is trying to produce an effect in
them by their recognition of his intention to that effect. This is a fairly
"complex" intention to ascribe to a being with no linguistic means of ful-
filling it. Does not Strawson fully embrace the feared absurdity when he
credits utterers and hearers with no prior conventions of communication
with being able to use their mutual "knowledge" as a "reason" for

"interpreting" utterances in the same way as before? While there are surely simple forms of communication that work along the broad lines sketched above, it is a stretch to couch them in such rationalistic terms (cf. Avramides 1989, 162-163). If we want to adopt a Gricean approach, communication is a deeply rational endeavor which is not instantiated by regularities among stimuli responses. Strawson's approach is just the first step in a long story about how we could get from stimuli to self-referential communication-intentions. Whatever happens when animals without language communicate, they do not communicate by wanting to get each other to understand or believe something in virtue of the recognition of that intention.[23]

The reason why the full force of the tension I have discussed has not been appreciated, seems to be that it is usually treated as a charge of inconsistency. Neale and Strawson focus on the question whether Grice can be reinterpreted consistently, that is whether his account rests on premises that undermine the project.[24] But the issue is not just whether there is a way of reinterpreting Grice consistently, but whether Grice's project can be reconstructed consistently while preserving its apparent aim. To my mind, it seems clear that Grice is interested in reducing meaning to intentions without semantic remainder rather than in merely explicating it in terms of the latter. One should keep in mind that timeless meaning, which is needed in any form of communication (even in nonliteral speech), is a function of a community of speakers intending it to mean what it means, which function

---

[23]   It could be objected that this artificially creates a problem for Grice because I am here sticking to the self-reflexive intention which has in later works been dropped both by Grice himself and by most philosophers drawing on Grice to further the project in their own ways. Were it not for the self-reflexive intention, the problem would not seem to persist. Having said that, my paper is intended to problematize the original account as presented in "Meaning" and is thus orthogonal to the issue of, say, the late Grice/Schiffer amendments. My aim has been to show that Searle's calls to save the Gricean account by returning to "Meaning" and its self-reflexive intention are to be rejected.

[24]   Neale's main concern in this regard is to show that Grice can a) account for the way in which conventional meaning plays a role in working out communicative intentions (cf. section 1 of this paper), and b) explain how the meaning of a sentence is (partly) determined by its parts. For these reasons, Neale rejects the view that Grice's account is either circular or absurd. See Neale (1992, 544, 550-552).

is itself a function of the ability of intending utterances to mean something. In this sense, Grice's account is not really an example of an intention-based semantics, as intentions must be conceived as carriers rather than constituents of meaning. Grice's original account can explain why utterances mean what they mean (rather than meaning something else), but it cannot explain how it comes about that we can even mean things by sharing utterances.

3. Conclusion

The analysis of Searle's American soldier example was meant to bring out, from the perspective of Grice's original account from "Meaning", the need and simultaneous inadequacy of having utterances rely on a shared repository of conventional meaning. If we postulate the reliance of intentions on prior meaning, we give up any hope of reducing timeless meaning to speaker's meaning without semantic remainder, and with it the project of an intention-based semantics. If we do not, we cannot explain how we can come to form or communicate complex intentions precisely because we have no rules or conventions to fall back on. The project of "Meaning" thus fails as an attempt at constructing an intention-based semantics: it is not possible to analyze timeless meaning in terms of speaker's meaning without semantic remainder, as speaker's meaning is itself reliant on prior timeless meaning. As I have tried to show, this is a serious blow to any attempts (even non-reductionist ones) at rehabilitating Grice's original account, which does indeed seem committed both to self-reflexive intentions and full analyzability of timeless meaning in terms of speaker's meaning.

References

Armstrong, D. M. (1971): Meaning and Communication. *The Philosophical Review* 80(4), 427-447.

Avramides, A. (1989): *Meaning and Mind. And Examination of a Gricean Account of Language*. Cambridge MA: MIT Press.

Bennett, J. (1973): The Meaning-Nominalist Strategy. *Foundations of Language* 10(1), 141-168.

Brandom, R. B. (1994): *Making It Explicit. Reasoning, Representing, and Discursive Commitment*. Cambridge MA: Harvard University Press.

Burge, T. (1979): Individualism and the Mental. *Midwest Studies in Philosophy* 4, 73-121.

Grandy, R. & Warner, R. (2017): Paul Grice. In: Zalta, E. N. (ed.): *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), <https://plato.stanford.edu/archives/win2017/entries/grice/>.

Grice, P. (1989a): Logic and Conversation. In: Grice, P.: *Studies in the Way of Words*. Cambridge MA: Harvard University Press, 22-40.

Grice, P. (1989b): Meaning. In: Grice, P.: *Studies in the Way of Words*. Cambridge MA: Harvard University Press, 213-223.

Grice, P. (1989c): Meaning Revisited. In: Grice, P.: *Studies in the Way of Words*. Cambridge MA: Harvard University Press, 283-303.

Grice, P. (1989d): Retrospective Epilogue. In: Grice, P.: *Studies in the Way of Words*. Cambridge MA: Harvard University Press, 339-385.

Grice, P. (1989e): Utterer's Meaning, Sentence-Meaning, and Word-Meaning. In: Grice, P.: *Studies in the Way of Words*. Cambridge MA: Harvard University Press, 117-137.

Grice, P. (1989f): Utterer's Meaning and Intentions. In: Grice, P.: *Studies in the Way of Words*. Cambridge MA: Harvard University Press, 86-116.

Kripke, S. A. (1982): *Wittgenstein on Rules and Private Language*. Cambridge MA: Harvard University Press.

Lycan, W. G. (2008): *Philosophy of Language. A Contemporary Introduction*. 2. Ed. NY/London: Routledge.

Martinich, A. P. (1984): *Communication and Reference*. Berlin/NY: Walter de Gruyter.

Neale, S. (1992): Paul Grice and the Philosophy of Language. *Linguistics and Philosophy* 15(5), 509-559.

Recanati, F. (1986): On Defining Communicative Intentions. *Mind & Language* 1(3), 213-242.

Schiffer, S. (1972): *Meaning*. Oxford: Clarendon Press.

Searle, J. R. (1969): *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Searle, J. (2007): Grice on Meaning: 50 Years Later. *Teorema* 36(2), 9-18.

SPERBER, D. & WILSON, D. (1986): *Relevance. Communication and Cognition*. Cambridge MA: Harvard University Press.

STRAWSON, P. F. (1971): *Meaning and Truth*. Reprinted in: P. F. Strawson: *Logico-Linguistic Papers*. London: Methuen, 170-189.

VLACH, F. (1981): Speaker's Meaning. *Linguistics and Philosophy* 4(3), 359-391.

YU, P. (1979): On the Gricean Program about Meaning. *Linguistics and Philosophy* 3, 273-288.

ZIFF, P. (1967): On H. P. Grice's Account of Meaning. *Analysis* 28, 1-8.