Ratio (new series) XXVIII 4 December 2015 0034-0006

PRAGMATICS AND PROCESSING

Bart Geurts and Paula Rubio-Fernández

Abstract
Gricean pragmatics has often been criticised for being implausible from a psychological point of view. This line of criticism is never backed up by empirical evidence, but more importantly, it ignores the fact that Grice never meant to advance a processing theory, in the first place. Taking our lead from Marr (1982), we distinguish between two levels of explanation: at the W-level, we are concerned with what agents do and why; at the H-level, we ask how agents do whatever it is they do. Whereas pragmatics is pitched at the W-level, processing theories are at the H-level. This is not to say that pragmatics has no implications for psychology at all, but it is to say that its implications are less direct than is often supposed.

1. Introduction

Human behaviour is often directed towards the achievement of goals. If the behaviour is collaborative, the goals are joint goals, and to the extent that linguistic communication is a form of collaboration, its goals are shared between interlocutors. Thanks to Grice, we have come to appreciate that viewing discourse as a cooperative enterprise pays explanatory dividends, though it must be noted that the Gricean approach has its limits, if only because speakers aren't always fully cooperative.

Grice's (1975a) theory of linguistic cooperation is cast in a somewhat peculiar mould. It takes the form of an overarching "Cooperative Principle" that enjoins speakers to design their utterances "such as is required" by the discourse goals. This principle is fleshed out by a set of "conversational maxims" which stipulate that one should try to speak relevantly, truthfully, non-redundantly, and so on. The presumption of cooperativity creates commitments on the part of the speaker. For example, an utterance of the sentence, "It's sunny", will normally commit the speaker to the belief that it is sunny, courtesy of the first Quality maxim. The audience's obverse of the speaker's commitment is that they are entitled to infer that he believes it is sunny.

In the post-Gricean literature, such inferences are usually called "conversational implicatures". Grice himself reserved this term for speakers' commitments, but as speakers' commitments and hearers' entitlements are two sides of the same coin, there is no issue here.

The Gricean view entails that, in order to analyse and explain the speaker's commitments, we can adopt the audience's point of view and imagine how they might reverse-engineer the speaker's utterance. Grice (1975a: 31) proposes the following schema for this procedure:

(1) – He has said that *p*;
– there is no reason to suppose that he is not observing the maxims, or at least the Cooperative Principle;
– he could not be doing this unless he thought that *q*;
– he knows (and knows that I know that he knows) that I can see that the supposition that he thinks that *q* is required;
– he has done nothing to stop me thinking that *q*;
– he intends me to think, or is at least willing to allow me to think, that *q*;
– and so he has implicated that *q*.

In the following exchange, for instance, this pattern of reasoning explains how Greengrocer's statement comes to implicate that he doesn't have papayas:

(2) Customer: Do you have papayas?
Greengrocer: We have very nice cantaloupes.

In his writings on pragmatics, Grice confined his attention to one particular type of illocutionary act, namely assertions, and his choice of maxims mirrors this limitation. In particular, the Quality maxims, which urge the speaker to be truthful and have adequate evidence for his utterances, are obviously restricted in their application. However, it is equally obvious that Gricean pragmatics extends not only to other illocutionary acts, but also to such linguistic acts as the production and interpretation of words, grammatical constructions, and intonation contours (cf. Geurts 2010: 182–187). To illustrate this point, consider the word "chestnut", which according to the OED has three established senses: it can be used to refer to (*i*) a glossy hard brown edible nut, (*ii*) the

large European tree that produces the edible chestnut, or (*iii*) a horse of a reddish-brown or yellowish-brown colour. Now suppose the following sentence is uttered:

(3) The chestnuts are shedding their leaves.

This utterance will naturally give rise to the inference that the speaker is using "chestnut" in the second sense. What justifies this inference? In a nutshell, the answer is that the inference is justified because, in the context of (3), a cooperative speaker would not use the word in either of the other senses. Spelled out in more detail, following Grice's schema:

(4) – She has used the word "chestnut";
– there is no reason to suppose that she is not observing the Cooperative Principle;
– she could not be doing this unless she intended to refer to trees of the genus Castanea, for this is one of the standard meanings of the word, and it fits our discourse purposes better than any of the others;
– she knows (and knows that I know that she knows) that I can see that the supposition that she intends the word in this meaning is required;
– she has done nothing to stop me thinking that this is so;
– she intends me to think, or is at least willing to allow me to think, that she intended to refer this type of tree;
– and so this is what she has implied.[1]

This may seem like an overworked use of Gricean reasoning, since we tend to take it for granted that speakers use words in conventional ways, but a moment's reflection will suffice to see that this is not always true: speakers may be mistaken about conventional word meanings, for example, or they may use words in figurative senses. Moreover, even if words were invariably used in a conventional way, this would still be a remarkable fact calling for an explanation: how does a linguistic community

[1] It is a matter of taste whether it would be wise to call such lexical inferences "conversational implicatures". Geurts (2010: 25) cautions against this on the grounds that that label has become firmly associated with illocutionary acts. However, this is merely a terminological issue.

converge on conventional word meanings (or any other conventions, for that matter) that are more or less stable? Lewis (1969, 1975) has shown how such questions can be addressed in a Gricean framework.

The style of reasoning exemplified in (1) and (4) is characteristic of Grice's own work and of Gricean pragmatics more generally. Its purpose is to make explicit why the hearer is *entitled* to draw certain inferences from the speaker's utterance, or what comes down to the same thing, why the speaker *commits* himself in certain ways by speaking as he does. These protracted trains of thought are hypothetical; they merely serve to unveil the pragmatic logic of a linguistic act. This core feature of the Gricean project, which is our main topic, has been missed by many authors, and even those who got the point were often puzzled, as the following passage from Warner (2001) illustrates:

> The problem, of course, is that people hardly ever reason this way when communicating. [. . .] So, what is the relation between the reasoning you *might* have engaged in and your understanding the sentence? How is there any explanatory power in the fact that, although you reached your understanding of the sentence *in some other way*, you *might* have reasoned your way to such an understanding? [. . .] This facet of his methodology was often noted and discussed at Berkeley during the 1970s. The worry was, of course, that people do not really reason explicitly in the way Grice would imagine. Once, when he was taken to task for this, he replied (with some exasperation), "But there must be a rational explanation!" (Warner 2001: x)

In this paper, we try to allay Warner's worry by showing how pragmatics, as understood by Grice, fits into a broader framework for explaining communication, which also includes a theory of processing.

2. "Such lengthy chains of inference"

As we have just seen, Gricean pragmatics is not concerned with the psychological processes that run their course while their owners produce and interpret words and sentences. Thus construed, pragmatics would be a patently silly enterprise. It comes as something of

a surprise, therefore, that this is precisely how Grice has been understood. In fact, one of the standard objections against the Gricean approach is that it makes communication impossibly hard from a psychological point of view (see Moore 2014 for a recent survey of criticism along these lines). For example, consider Wilson's (2000) remarks a propos of Grice's general schema for deriving conversational implicatures, cited in (1):

> Grice seems to have thought of the attribution of meaning as involving a form of conscious, discursive reasoning [. . .] It is hard to imagine even adults going through such lengthy chains of inference in the attribution of speaker meanings. (Wilson 2000: 415–416)

The same objection has been made by Origgi and Sperber (2000) and Sperber and Wilson (2002), among others. Wilson's passage contains three common misconceptions:

i. Gricean pragmatics adopts a mentalistic stance in the sense that it is concerned with the internal states and processes underlying interpretation.
ii. These states and processes are available to consciousness.
iii. It's all way too complicated: in reality, interpretative processes are a lot simpler than Grice would have us believe.

These claims are linked: (*ii*) and (*iii*) presuppose (*i*), and (*iii*) is based on (*ii*), if only tacitly: charges like (*iii*) are never backed up by experimental evidence, and therefore we have to suppose that they are based on introspection.

Millikan (1984) raises a similar issue about Grice's claim that communication involves, on the part of the speaker, intentions about intentions, etc., as well as the recognition of all these higher-order intentions, on the part of the hearer:

> The genuine communicator is not a creature that, in the process of every speech act, intends that his hearer believe that he intends him to – etc. How inefficient that would be, if we take having intentions and beliefs to be real modifications of the nervous system! (Millikan 1984: 69)

Millikan's charge of "inefficiency" is not supported by empirical evidence, so her argument, too, is based on introspection, and

her complaint is the same as Wilson's: the Gricean account is just too complex to be psychologically plausible. Again, it is taken for granted that Grice's take on communication is mentalistic.[2]

As discussed in the introduction, and at greater length by Warner (2001) and Saul (2002), among others, the higher-order intentions and "lengthy chains of inference" that pervade the Gricean literature are not to be understood as hypotheses about ordinary mental states and processes, and Grice himself, too, is clear about this. However, we are not so much concerned with the correct exegesis of Grice's corpus as with the fact that neither the merits nor the demerits of Gricean pragmatics are contingent on a mentalistic construal of intentions and other attitudes. Although the vocabulary of Gricean pragmatics is unabashedly psychological in its reliance on propositional-attitude concepts (intentions, beliefs, desires, and the like), it is neutral on the question of whether propositional attitudes are to be viewed as (internal) mental states and processes.

To drive home this point, suppose that we were all logical behaviourists; suppose, that is, that we all defined our attitude concepts in terms of behavioural patterns and dispositions.[3] So we're imagining that (5a) is a shorthand for something like (5b):

(5) a. Agnes intends to eat a banana.
    b. Agnes is prone to exhibit the sort of behaviour that will increase the likelihood of her obtaining the opportunity to eat a banana.

(One sees why such abbreviations might be useful.) This line of analysis is generally agreed to be problematic, but that is as it may be, for we are making a conceptual point, not an empirical

[2] There is a more general tendency to take it for granted that any propositional-attitude talk is necessarily mentalistic. It can also be seen in the recent philosophy of mind, for instance, where theories of "mindreading" have prompted caricatures and criticisms similar to those fielded by Wilson and Millikan, and often expressed in similar terms: "our understanding of others chiefly engages detached intellectual processes, moving by inference from one belief to the other." (Zahavi 2007: 26) For more examples, see the papers collected by Hutto and Ratcliffe (2007). Spaulding (2010) provides an excellent overview of, and responses to, various objections in this vein.

[3] We note in passing that although Grice was certainly no behaviourist, he did emphasise that "attributions of psychological states owe any claim to truth they may have to their potentialities for the explanation of behaviour." Grice (1975b: 157)

one, and the point is just this: if all of us were logical behaviourists, Grice's schema for deriving conversational implicatures would not be affected in any way, since it doesn't rely on the assumption that propositional attitudes are mental states; the same goes for Grice's view that communication essentially involves higher-order intentions.

If Gricean pragmatics is not mentalistic, then *a fortiori* it is not about conscious mental states or processes, nor is there any reason to suppose that its alleged complexity has any bearing on the complexity of mental states and processes. Hence, all of Wilson's and Millikan's claims about Gricean pragmatics go by the board.

Still, it is not unreasonable to ask, *what if* Wilson and Millikan had been right that Gricean pragmatics is mentalistic? First, would there then be any reason to assume that, on a Gricean account, pragmatic processes are conscious? As far as we can tell, there is no reason to suppose that this is so, nor has it ever been argued for. Consciousness is a red herring.

Secondly, if pragmatic processes are not, or not always, open to consciousness, introspective judgments on the complexity of such processes must be taken with a grain of salt. Moreover, there are good reasons to believe that "phenomenological evidence" on mental processes is generally unreliable (cf. Carruthers 2009, Spaulding 2010). Consider visual perception. What could be easier than seeing a chair, for example? Anyone with normal or corrected-to-normal vision can do it; even young children do it with ease. Nevertheless, all textbooks agree that the mental processes underlying visual perception are bafflingly intricate, and the complexity of Gricean reasoning pales when compared to that of seeing a chair. If we are to gauge the complexity of mental processes, introspective evidence is as good as no evidence, and more misleading.

In sum, if Gricean pragmatics was a processing theory, any claim to the effect that it is too complex would have to be based on solid empirical or analytical evidence, which thus far has not been forthcoming. So even if we grant Wilson and Millikan the premiss that those lengthy chains of reasoning and higher-order intentions are meant to describe mental processes and/or representations, their objections are void. But at the end of the day, none of this matters, since their premiss is false.

3. What and how

If the Gricean project is not about processing, then what is it about? In order to answer this question, we appropriate Marr's (1982) general framework for describing computational systems, though we will substitute our own terminology. Marr proposes a three-way distinction between computational, algorithmic/representational, and physical levels. We adopt the same division, but feel that Marr's labels for the first two levels are potentially confusing, and therefore will speak of "W-level" and "H-level", respectively, where W stands for "what" and "why" and H for "how"; Marr's physical level will not figure in our story under any name.

We will introduce the W/H distinction with the help of an example discussed by Marr: that of a cash register. When viewed at the W-level, we are concerned with *what* a cash register does, and *why*. The what question is answered by elementary number theory: addition is an operation on pairs of numbers, whose properties are well understood: it is commutative: $a + b = b + a$; it is associative: $(a + b) + c = a + (b + c)$; there is a zero element, which adds nothing: $a + 0 = a$; and so on.

Marr's answer to the why question is that we "intuitively feel" that the appropriate rules for combining prices are such that, taken together, they require addition. For instance, the order in which you present your goods to the cashier should not affect the total price, which is why price calculation must be commutative. Things are a bit more complicated than this, if only because supermarkets lure customers with all manner of discount schemes, like "Buy one, get one for free!". Apparently, there is an evolutionary dynamic with sellers trying out various strategies whose fitness is determined by buyers' selective pressures, resulting in mixed strategies in which addition is the principal operation for combining prices, though not the only one.

Whereas a W-level theory describes a system from the outside, so to speak, an H-level theory deals in internal processes, states, and representations. In the case of the cash register this involves representations for numbers and an algorithm that maps input representations to output representations, in a way that is sufficiently faithful to the laws of addition. There are many different representations of numbers and many different algorithms for addition, and choices will be dictated, inter alia, by considerations of efficiency and accuracy.

W-level and H-level theories constrain but don't determine one another. If we know that the purpose of a device is to add numbers, this restricts the range of plausible hypotheses about its internal workings; we wouldn't expect to find an algorithm for playing noughts and crosses there, for instance. Conversely, if it was somehow given that a device implemented a reasonably reliable algorithm for adding numbers, it would be a fair guess that adding numbers was one of its purposes. Note, however, that in general the W-level precedes the H-level, both in the logic of discovery and the logic of justification. Normally speaking, the hypothesis that the purpose of a device is addition would precede the hypothesis that the device performs additions by means of this or that algorithm; and normally speaking, the reason why a device implements an addition algorithm is that calculating sums is one of its purposes.

In general, the systems revealed by a process analysis will fail to provide a completely faithful implementation of the W-level theory. This holds true even in the case of the cash register. For one thing, there is always a limit to the size of the numbers that can be manipulated, even if that limit is astronomical. For another, decimal numbers are usually converted into binary ones, and in general, the decimal numbers entered by the cashier are merely approximated by their binary counterparts. For supermarkets and their customers, these limitations are of little concern, because the numbers they deal with are relatively small, and such inaccuracies as do occur are rarely significant. But in many other cases, the discrepancies between the W- and H-levels do matter. For example, when addition is performed by humans rather than electronic devices, the likelihood of error will increase dramatically.

It is sometimes said that a significant amount of slack between the two levels shows that a W-level theory wasn't needed in the first place. It is a well-established fact, for example, that people perform poorly on various reasoning tasks, which has caused quite a few scholars to argue that logic is irrelevant to our understanding of human reasoning (e.g., Evans 2002). This is a mistake not only because it is informed by a parochial view on logic, but also because an adequate understanding of what is done and why does not entail that it is done well. No doubt the logical prowess of our species has its limits, and in some respects they may be severe, but that does nothing to alter the fact that, for a great many purposes, we are designed to reason logically.

More generally, we wish to emphasise, echoing Marr, that the W-level always deserves serious theoretical scrutiny; which it often doesn't get. Addition may seem a simple thing, but the general theory of addition is not that simple, and as we have seen, the exact role that addition is to play in a comprehensive account of cash registers is not cut and dried either. So even in this case, a W-level analysis faces non-trivial issues, and it is only to be expected that a W-level theory of human behaviour will be vastly more complex. This is an important point, because there is a widespread tendency, in the social sciences as well as the humanities, to treat W-level analyses perfunctorily, if at all, and head straight for the processes and representations. This tendency may be partly due to a penchant for mentalistic explanations, but another part may be that scholars often fail to distinguish between levels that must be kept apart, no matter how we choose to label them. This failure to keep one's levels straight has given rise to a fair amount of confusion; the misconceptions about Gricean pragmatics are a case in point, but there are many more.

Admittedly, it is not always easy to distinguish between W and H, because the same theory may sometimes be located on either level, or both. Let's explain this with the help of an example. Suppose we are concerned with human reasoning, specifically with the way people deal with inferences like the following:

(6) If it is raining or snowing, and there is no reason why we should go out, then we should stay in.
It is snowing.
There is no reason why we should go out.
Ergo: We should stay in.

Assume, if only for expository convenience, that classical propositional logic is the right W-level theory for this task. Then there are two very different ways of spelling out that theory:

SEMANTIC: We lay down the truth conditions for sentences of the form "*p* and *q*", "if *p* then *q*", etc., and define the notion of entailment as follows:

q is *entailed* by a set of premisses $p_1, \ldots, p_n$ iff *q* is true whenever $p_1, \ldots, p_n$ are true.

PROOF-THEORETIC: We lay down a set of proof rules, each of which permits us to derive a conclusion from a set of premisses. For example:

– If p and q are given, then "p and q" is derivable.
– If either p or q are given, then "p or q" is derivable.
– If p and "if p then q" are given, then q is derivable.
– . . .

Then we define the notion of provability as follows:
q is *provable* from a set of premisses $p_1, \ldots, p_n$ iff q can be obtained from $p_1, \ldots, p_n$ by applying the derivation rules.

Since propositional logic is sound and complete, these two consequence relations, entailment and provability, are equivalent in the following sense:

q is entailed by $p_1, \ldots, p_n$ iff q is provable from $p_1, \ldots, p_n$

It follows that if we use entailment in our W-level theory of propositional inference, we might as well use provability, and vice versa. Concretely, if we want to describe *what* Betty was doing when she inferred (6d) from (6a–c), it doesn't make any difference whether we say that Betty decided that (6d) is entailed by (6a–c) or that she decided that (6d) is provable from (6a–c).

However, despite the fact that the semantic and proof-theoretic approaches are interchangeable at the W-level, they part ways at the H-level. For, whereas the semantic approach leaves us without a clue as to how a conclusion might be drawn from a set of premisses, a proof-theoretic system may be incorporated in a processing model of human reasoning. In fact, one of the more popular ideas in the psychology of reasoning has always been that people actually use proof rules like the ones listed above (see Rips 1994 for a detailed proposal). Put otherwise, whereas the notion of entailment is unlikely to play any part in a processing model of reasoning, the notion of provability might well be included in such a model.

This little case study yields two important observations. First, theories that are very different may turn out to be equivalent at the W-level. Secondly, it is possible for the same theory to serve at the W- and H-levels. The first observation serves at a warning that differences between theories do not always matter, regardless how

substantial they may seem to be. The second observation may go some way to explain why the two levels are often mixed up. In particular, it may help to explain why Gricean pragmatics has so often been misconstrued.

4. Putting pragmatics in its place

At this point, it shouldn't come as a surprise that we propose to view Gricean pragmatics as a W-level theory, which is about what speakers do and why. Take, for example, B's contribution to the following exchange:

(7) A: What's the time?
B: Half past three.

What happened? B expressed his belief that the time was 15:30, intending A to know that this was so. Why? A had indicated that she wanted to know the time, and B's response showed that he accepted this as a common goal, which his utterance was intended to achieve. In the same spirit, if A deploys the word "monosemic" in the context of (8),

(8) "Polysemic" is monosemic.

then what A does is expressing the property of having only one meaning, and the reason why she does so is that this property is to serve as a building block for the assertion that "polysemic" is a monosemic word.

The broad picture is this. Gricean pragmatics aims to provide a W-level account of communication, analysing speakers' and hearers' behaviour in terms of their propositional attitudes (what) and communicative goals (why) on the assumption that, by and large, speakers try to be cooperative. The chief corollaries of such an analysis are conversational implicatures, which may be seen either as commitments incurred by the speaker, or as the corresponding inferences the hearer is entitled to draw.

The Gricean approach is neutral with respect to the ontological status of propositional attitudes; specifically, it is not mentalistic. Since communication is a form of social interaction, studying pragmatics on the premiss that it is all "in the mind" is like trying to understand traffic by studying car engines. But this is not to

deny the importance of investigating the mental states and processes that enable communication, nor that of elucidating the connections between pragmatics and pragmatic minds.

Though it is not unreasonable to expect that, to some degree at least, a process model will mirror the corresponding W-level description, we have seen that this need not be the case. There are indefinitely many ways of doing addition, good and poor, fast and slow, and although some of these may be similar to our preferred W-level analysis of addition, there is no reason to suppose that those processing models showing the highest resemblance must be any better than the others. The same holds for the relation between pragmatics and processing. This point is important enough to deserve an example. Consider a satellite navigation system that produces turn-by-turn directions: "After 80 metres, turn left", "Go straight at the roundabout, second exit", "Try to make a U-turn", etc. Such systems interact with their users in fairly complex ways, which it is natural to interpret in Gricean terms:

- Driver sets a common goal by choosing a destination.
- Satnav provides instructions to guide Driver.
- Driver tries to follow these instructions, continually modifying the car's GPS coordinates while doing so.
- Satnav's instructions are designed to be cooperative; i.e., to be correct, relevant, non-redundant, and so on.
- Satnav's instructions license implicatures. For example, if Satnav doesn't tell Driver to turn left at the upcoming intersection, then Driver is entitled to infer that she shouldn't do so.

In short, Driver and Satnav engage in a non-trivial form of cooperative interaction. But when we turn to Satnav's internal processes and representations, we are in for a disappointment: very likely, its production algorithm makes do with canned phrases triggered by a handful of parameter settings. At the H-level, Satnav is a mindless gadget.

One might be tempted to object that such a simple device couldn't be genuinely goal-directed, responsive to the driver's actions, cooperative, and so on. However, this objection involves a mixing of levels. To say that Satnav's functioning couldn't be any of these things because it lacks the procedural sophistication to match is like claiming that backgammon must be a stupid

game because all the pieces are identical. "But surely Satnav doesn't *literally* intend to send Driver this or that way?" That depends on what is meant by the verb "intend". Surely, its everyday, non-technical meaning is not sufficiently stable and precise to settle the matter, and technical meanings are stipulative, so depending on one's purposes the answer may come out as yes, no, or neither.

It is not unusual for complex patterns of behaviour to emerge from simple mechanical principles. A well-known case in point is swarming behaviour in birds and other animals. The shape and movement of a starling flock, for instance, is determined by each bird maintaining its position relative to the six or seven animals closest to it (Ballerini et al. 2008). The same goes, mutatis mutandis, for phenomena closer to the ground, like traffic flow (Orosz et al. 2009) or catching a ball while running (McLeod and Dienes 1996, Kistemaker et al. 2009). These and many other examples prove that, no matter how intricate a form of behaviour may appear to be, it need not be caused by underlying mechanisms of commensurate complexity.

Nevertheless, it seems not unreasonable to expect at least some similarities between pragmatics and processing, if only because, occasionally, speakers are overtly Gricean ("If you knew, why didn't you say so?"). In particular, since Gricean pragmatics essentially concerns speakers' and hearers' propositional attitudes, we should expect that attitude representations play at least some role in processing. In the following section, we argue that that role is likely to be considerable.

5. Processing attitudes

Much of our mental lives is routine, and it seems unlikely that we stop at every word to ask ourselves what the speaker's intentions are. What if it were the case that much or even most of the processing underwriting communication was automatic? According to Millikan (1984), it would show that, for the most part, communication is not Gricean:

> The truth in Grice's model is that we have the ability to interrupt and prevent the automatic running on of our talking and our doing-and-believing-what-we-are-told equipment, and assume others have this ability too. [. . .] The true

> communicator is in a position to tinker with the mechanisms of normal language flow, is sensitive to the symptoms that the other is tinkering with these mechanisms, and can rise above these automatic mechanisms if necessary. (Millikan 1984: 69)

Although Millikan's argument is somewhat opaque, its main thrust is clear enough. Most of the time, language users are on autopilot: in the normal run of events, speaking and understanding are "automatic". Millikan (1984, 2005) even goes so far as to classify language understanding as a form of perception. Still, every now and then, it is necessary to "rise above these automatic mechanisms", and it is then, and only then, that we become true Griceans; that is to say, only then do we start considering each other's beliefs, intentions, and so on. "Gricean processing", thus understood, is a fallback option, which is resorted to only when the going gets tough;[4] or as Pickering and Garrod (2004: 180) put it, "normal conversation does not routinely require modeling the interlocutor's mind". The same idea has been floated by Gauker (2003) and Apperly (2011), among others.

There are lots of things to be said about this line of argument, but we will confine ourselves to three. First, Millikan offers no evidence whatsoever for her sweeping statement that "the mechanisms of normal language flow" are automatic. Apparently, Millikan's conviction is based on introspection, and as we argued in §2, introspection is a poor guide to the true nature of mental processes. Furthermore, there is no experimental evidence for Millikan's view, either (and we're writing three decades after she first went on record with it). In short, Millikan's claim is just a claim.[5]

Secondly, the suggestion that communication is mostly automatic is plainly outrageous when taken literally: using language isn't anything like a knee-jerk reflex. But if that claim is to be taken non-literally, it is something of a mystery what Millikan's position comes to. However, since Millikan is at pains to argue that speakers' intentions and other atttitudes are not essential to

[4] Since "Gricean" is a W-level notion, we enclose "Gricean processing" in scare quotes to emphasise that this is a derived and non-literal use of the word.

[5] Note, incidentally, that the fast/slow (or automatic/controlled) distinction cross-cuts the implicit/explicit distinction (Carruthers 2013).

communication, it would seem that her use of the term "automatic" is meant to exclude, at the very least, mental processes in the speaker and hearer which take such attitudes into account, and that is how we will interpret her view. Thus construed, Millikan's view is amply discredited by empirical evidence.

These days, propositional attitudes seem to have gained an undeserved reputation for being inherently complex. The alleged complexity is variously blamed on the fact that they are unobservable; that reading them involves abductive reasoning (aka "inference to the best explanation"); that propositional attitudes have a normative aspect; and so on (see, e.g., Apperly and Butterfill 2009). However, even if dealing with propositional attitudes can be hard, we're not aware of any evidence that it is necessarily so. On the contrary, it seems obvious that our everyday social interactions are steeped in suppositions about each other's beliefs, desires, intentions, and so on, and since much of our social interaction is quite smooth, it can't be true that processing information about propositional attitudes is always hard.

In some parts of philosophy and psychology, it has become fashionable to deny this. Millikan was an early representative of the revisionist undercurrent, but more recently this line has attracted such unlikely bedfellows as Gauker (2003), Keysar et al. (2003), Pickering and Garrod (2004), Zahavi (2007), and Apperly (2011), among many others. The leitmotif in all these various discussions is the same: Yes, people are perfectly capable of attributing propositional attitudes to one another; but no, this ability does not play a central role in everyday social interaction.

To the best of our knowledge, there is no compelling evidence to back up this view. At best, there are isolated findings suggesting that our practice of attitude attribution isn't perfect; which was to be expected. But there is no serious body of data to support the claim that the attribution of beliefs and other attitudes is a rare thing. Contrariwise, there is solid evidence for the opposite view. Communication would be a complete mess or worse if speakers failed to routinely take into account their audiences' psychological states. If someone tells a colleague that the photocopier has broken down, that their firm is bankrupt, or that she likes his tie, she normally does so because she assumes that he didn't know and/or has an interest in knowing. On the other hand, if she *doesn't* tell him that eggs are a good source of protein, that two is even, or that he is wearing trousers, it must be because she assumes that he knows all this. In short, speakers routinely tailor

their messages to what they take to be their audience's knowledge, needs, and preferences. Communication would break down if they didn't do so.

Based on the foregoing considerations, we take it to be obvious that language production relies heavily on "reading hearers' minds". It is less obvious that the converse is also true; that is to say, that language understanding requires that the hearer constantly attend to the speaker's propositional attitudes. However, we believe that there is more than enough evidence to show that this holds as well, but since a full-scale discussion of the data is out of the question here, we will go through a handful of representative cases only.

□ While riding on a tram, an old lady you never met before turns to you and asks:

(10) Isn't it a shame that our prime minister still isn't married?

Not a common way of opening a conversation, by any means, but now consider an alternative scenario, in which the same lady asks:

(11) Isn't it a shame that your brother still isn't married?

Very likely, your first thought will be something like this: "How on earth could she know that I have a brother and that he is single?" An utterance of (10) signals that it is common ground that there is a prime minister and that he is single; likewise, an utterance (11) signals that it is common ground that the addressee has a brother and that *he* is single. Whereas the first pair of presuppositions can safely be taken for granted among fellow countrymen, the second pair requires knowledge that is less widely available. The contrast between (10) and (11) shows that hearers *check* whether speakers can plausibly possess the knowledge which they present as part of the common ground.

□ Suppose Mr. Chekhov points at an anthology of his short stories, and utters (12a):

(12) a. Most of these stories were written for popular magazines.

b. All of these stories were written for popular magazines.

Chekhov's utterance would clearly license the inference that (12b) is false. Now imagine that Mr. Bean points at the same book and utters the same sentence, i.e. (12a). In this case,

the hearer will be less inclined to infer that (12b) is false, and more likely to conclude that the speaker doesn't know whether or not (12b) is true. This contrast isn't hard to explain: unlike Mr. Bean, Chekhov is expected to know whether or not (12b) is true. Apparently, the pragmatic inferences an utterance gives rise to are contingent on the hearer's assessment of the speaker's expertise.[6]

□ Gesturing towards an arrangement of flowers in various colours, somebody asks: "How would you describe the colour of this flower?" Clark et al. (1983) present evidence showing that, in order to determine the most likely referent in cases like this, hearers rely, inter alia, on assessments of speakers' goals and knowledge. In another experiment, participants were shown a picture of US president Ronald Reagan sitting with David Stockman, then the director of the Office of Management and Budget; the assumption being that, at the time, Reagan was much better known than Stockman. Each participant was asked one of the following questions:

(13) a. You know who this man is, don't you?
b. Do you have any idea at all who this man is?

In asking (13a), the interviewer presupposed that the intended referent was known, and therefore they inferred he must have Reagan in mind; when he presupposed the opposite, their preference shifted to Stockman.

□ Compare the following sentences:

(14) a. It's sunny, but Agnes doesn't know it.
b. It's sunny, but I don't know it.

Whereas (14a) is perfectly coherent and consistent, (14b) seems odd. This is remarkable, because the two sentences have very similar meanings. Indeed, if (14b) is uttered by Agnes, it has the same truth conditions as (14a). Still, it's not hard to see why an utterance of (14b) should be infelicitous: if Agnes asserts that it is sunny, the hearer is entitled to infer (courtesy of the first Quality maxim) that Agnes believes what she says, and provided he *makes* that inference, he will see that it contradicts the second half of Agnes's statement; whereas if the inference is not made, the utterance will go

[6] This type of case is a stock-in-trade of the Gricean literature on quantity implicatures. The now-standard analysis is Soames's (1982). See Geurts (2010: 37–42) for extensive discussion, and Goodman and Stuhlmüller (2013) for experimental evidence.

through without a hitch. Now it seems easy to see that (14b) is infelicitous, and experimental evidence confirms that its oddity is registered without delay, compared to various controls (Geurts and van Tiel 2015). *Pace* Millikan and others, this indicates that hearers standardly interpret an assertion of *p* to imply that the speaker believes *p* to be true.

Observations like these can be multiplied without difficulty, and they suggest quite forcefully that hearers *routinely* monitor speakers' beliefs, desires, and other attitudes; which is consistent with the Gricean view, but goes against the notion that Gricean reasoning is necessarily effortful and for special occasions only.

Let's take stock. It has been maintained by Millikan and others that, by and large, communication is an automatic affair, which is taken to imply that, in the normal run of events, speakers and hearers are not concerned with each other's propositional attitudes. In the foregoing we have argued against this view on empirical grounds: there is precious little evidence in its favour, and rather abundant evidence indicating the exact opposite.

To conclude this discussion, we turn to a key premiss underlying Millikan's position, namely that if it is the case that language production and interpretation are mostly discharged by automatic mechanisms, then the Gricean approach is mostly false. This assumption is problematic no matter how you look at it, and our satnav example points up one way in which Millikan's reasoning is incorrect: a communication system whose function is thoroughly Gricean may be be implemented, at the H-level, by a mindless routine. But Millikan doesn't acknowledge the W/H distinction, or any distinction between levels, for that matter. Her sole concern is with processing. So presumably we should rephrase her claim as follows: If the mental processes that underlie linguistic communication are mostly automatic while "Gricean processes" come into play only rarely, then that shows that the system is mostly non-Gricean.

Note that when interpreted thus (and it's the most charitable interpretation we can think of), Millikan's disagreement with Grice is purely verbal, for she is concerned with processing and he wasn't. Hence, as mentioned before, in this connection our use of the term "Gricean" is non-literal and merely for convenience. However, even with these provisos, Millikan's argument is a dubious one. To explain why, let's pretend that the mental processes dealing with language understanding form a self-

contained system and let's concede, if only for argument's sake, that this system operates in automatic mode most of the time, with Gricean processes entering the fray only if needed. Hence, the system will have at least two tasks to contend with, which for convenience we assign to two separate modules. SYSTEM 1 is a low-level module in charge of the legwork, whose modus operandi is fully automatic. SYSTEM 2 is in charge of Gricean interpretation; it is dormant most of the time, and it only wakes up when the going gets rough.

Two-system models of cognition have been popular for some time, and they have been applied to social cognition, too, notably by Apperly and his associates (Apperly 2011). Such models are attractive because they separate chores that can be dealt with routinely, subliminally, and at virtually no cost, from tasks that require attention, deliberation, and effort. However, two-system theories invariably give rise to the issue of how System 1 and System 2 interact. To bring out the problem, let's suppose that there is a SYSTEM 1.5, which monitors the proceedings, and especially System 1's output, and if necessary, passes control to System 2. The problem is simply this: How is System 1.5 to decide when to call on System 2 without having a great deal of the sophistication that is supposed to be System 2's prerogative? If it is the case that substantial power and resources are required just to determine whether System 2 must become active, the whole point of a two-system architecture is lost. We believe that pragmatic processing might well be such a case.

To explain, let's revisit the contrast between (14a) and (14b), repeated here for convenience:

(14) a. It's sunny, but Agnes doesn't know it.
b. It's sunny, but I don't know it.

Crucially, at the semantic level these sentences are not very different. Indeed, when uttered by Agnes, (14b) describes the same state of affairs as (14a). Nevertheless, unlike (14a), (14b) is infelicitous no matter who the speaker is, its infelicity is evident and readily perceived, which implies that the hearer draw the inference that the speaker believes that it is sunny. By hypothesis, that inference is for System 2 to make. But why should it be made in the first place? After all, absent that inference, (14b) is as consistent as (14a) is. Therefore, in order to detect that something is wrong and decide that System 2 must spring into action, System

1.5 has to engage in reasoning about the speaker's beliefs, which was supposed to be the province of System 2.

Other examples will serve to make the same point. Take the case of the old lady on the tram, for example. Why is it odd for her to ask, "Isn't it a shame that your brother still isn't married?" Obviously, it's because her utterance licenses the inference that she knows you have an unmarried brother, and provided you make that inference, you may ask yourself how she could know that. Again, in order to detect an anomaly, System 1.5 needs to assess the speaker's knowledge, which *ex hypothesi* is part of System 2's job description.

To sum up, we find no merit in the argument that the Gricean approach is mostly irrelevant because communication is largely routine. This is not to deny that many of the processes underwriting language understanding and production are automatic. However, it is to say that it seems unlikely that such processes form an autonomous system that operates more or less independently from higher-level processes, which become active under special circumstances only.

6. Conclusion

In the introduction to this paper, we cited Warner's introduction to Grice's *Apects of reason* (2001):

> So, what is the relation between the reasoning you *might* have engaged in and your understanding the sentence? How is there any explanatory power in the fact that, although you reached your understanding of the sentence *in some other way*, you *might* have reasoned your way to such an understanding? (Warner 2001: x)

How can a hypothetical chain of inferences resulting in an implicature contribute to our understanding of communication when it need not reflect the interlocutors' mental processes? We have tried to answer this question by making a distinction between two interdependent levels of description and explanation, W and H, and locating pragmatics on the W-level, simply on the grounds that communication is a public affair. After all, it takes place between people: speakers incur commitments by performing linguistic acts, thus entitling hearers to draw certain inferences.

That's *what* speakers and hearers do, and it's what pragmatics is about; this is the W-perspective on communication. The H-perspective is concerned with the mental processes allowing us to explain *how* speakers and hearers do whatever they do.

The two perspectives will have to mesh, but how they will mesh is very much an open question. We have seen that, in principle, a W-level theory and its H-level counterpart need not constrain each other very strongly, but we have also advocated the view that, in fact, pragmatics and processing are coupled relatively tightly. In particular, we have argued that propositional attitudes and their representations figure essentially on the W- and H-levels, respectively.

We maintain that propositional-attitude representations are routinely involved in the mental processing underlying linguistic communication. Our argument goes against a multi-source current in philosophy and psychology which seeks to downplay the importance of propositional attitudes. That current is fed by a great many sources, but there is one common bogeyman: attitude attribution has come to be viewed as a conscious, intellectual process which is inherently complex and effortful. To a large extent, this is what drives otherwise disparate attempts at "simplifying" communication by keeping propositional attitudes at a distance, e.g., by quarantining them in some System 2 or other. However, the stigmatisation of propositional attitudes is groundless and undeserved. There is no good reason for supposing that the attribution of beliefs, desires, and intentions is necessarily conscious, effortful, and slow, and there is quite compelling evidence for assuming the opposite.

Acknowledgements

We would like to thank Emma Borg and Nat Hansen for their very helpful comments on the first version of this paper. The authors were supported by grants from the Netherlands Organisation for Scientific Research (NWO) and the Research Council of Norway (project no. 230718), respectively.

References

Apperly, I. A. (2011). *Mindreaders: the cognitive basis of "theory of mind"*. Hove: Psychology Press.

Apperly, I. A. and S. A. Butterfill (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological review* 116: 953–970.

Ballerini, M., N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale and V. Zdravkovic (2008). Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study. *Proceedings of the National Academy of Sciences* 105: 1232–1237.

Carruthers, P. (2009). How we know our own minds: the relationship between mindreading and metacognition. *Behaviorial and brain sciences* 32: 121–182.

—— (2013). Mindreading in infancy. *Mind and language* 28: 141–172.

Clark, H. H., R. Schreuder and S. Buttrick (1983). Common ground and the understanding of demonstrative reference. *Journal of verbal learning and verbal behavior* 22: 245–258.

Evans, J. S. (2002). Logic and human reasoning: an assessment of the deduction paradigm. *Psychological bulletin* 128: 978–996.

Gauker, C. (2003). *Words without meaning*. Cambridge, Massachusetts: MIT Press.

Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.

Geurts, B. and B. van Tiel (2015). Automatic inferences about speakers' beliefs. In preparation.

Goodman, N. D. and A. Stuhlmüller (2013). Knowledge and implicature: modeling language understanding as social cognition. *Topics in cognitive science* 5: 173–184.

Grice, P. (1975a). Logic and conversation. In P. Cole and J. Morgan (Eds.), *Syntax and semantics 3: Speech acts*, pp. 41–58. New York: Academic Press. Reprinted in and cited from Grice (1989: 22–40).

—— (1975b). Method in philosophical psychology (from the banal to the bizarre). *Proceedings and addresses of the American Philosophical Association,* 48: 23–53. Reprinted in Grice (1991: 121–161).

—— (1989). *Studies in the way of words*. Cambridge, Massachusetts: Harvard University Press.

—— (1991). *The conception of value*. Oxford University Press.

—— (2001). *Aspects of reason* (edited by Richard Warner). Oxford University Press.

Hutto, D. D. and M. Ratcliffe (Eds.) (2007). *Folk psychology re-assessed*. Dordrecht: Springer.

Keysar, B., S. Lina and D. J. Barr (2003). Limits on theory of mind use in adults. *Cognition* 89: 25–41.

Kistemaker, D., H. Faber and P. Beek (2009). Catching fly balls: a simulation study of the Chapman strategy. *Human movement science* 28: 236–249.

Lewis, D. K. (1969). *Convention*. Cambridge, Massachusetts: Harvard University Press.

—— (1975). Languages and language. In *Minnesota studies in the philosophy of science*, pp. 3–35. Minnesota University Press.

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman and company.

McLeod, P. and Z. Dienes (1996). Do fielders know where to go to catch the ball or only how to get there? *Journal of experimental psychology: human perception and performance* 22: 531–543.

Millikan, R. G. (1984). *Language, thought, and other biological categories*. Cambridge, Massachusetts: MIT Press.

—— (2005). *Language: a biological model*. Oxford University Press.

Moore, R. (2014). Ontogenetic constraints on Grice's theory of communication. In D. Matthews (Ed.), *Pragmatic development in first language acquisition*, pp. 87–104. John Benjamins.

Origgi, G. and D. Sperber (2000). Evolution, communication and the proper function of language. In P. Carruthers and A. Chamberlain (Eds.), *Evolution and the human mind: modularity, language and meta-cognition*, pp. 140–169. Cambridge University Press.

Orosz, G., R. E. Wilson, R. Szalai and G. Stépán (2009). Exciting traffic jams: nonlinear phenomena behind traffic jam formation on highways. *Physical review E* 80(046205): 1–12.

Pickering, M. J. and S. Garrod (2004). Toward a mechanistic psychology of dialogue. *Behaviorial and brain sciences* 27: 169–190.
Rips, L. J. (1994). *The psychology of proof: deductive reasoning in human thinking*. Cambridge, Massachusetts: MIT Press.
Saul, J. M. (2002). What is said and psychological reality: Grice's project and Relevance theorists' criticisms. *Linguistics and philosophy* 25: 347–372.
Soames, S. (1982). How presuppositions are inherited: a solution to the projection problem. *Linguistic inquiry* 13: 483–545.
Spaulding, S. (2010). Embodied cognition and mindreading. *Mind and language* 25: 119–140.
Sperber, D. and D. Wilson (2002). Pragmatics, modularity and mind-reading. *Mind and language* 17: 3–23.
Warner, R. (2001). Introduction: Grice on reasons and rationality. In Grice (2001: vii–xxxviii).
Wilson, D. (2000). Metarepresentation in linguistic communication. In D. Sperber (Ed.), *Metarepresentations: a multidisciplinary perspective*, pp. 411–448. Oxford University Press.
Zahavi, D. (2007). Expression and empathy. In D. Hutto and M. Ratcliffe (Eds.), *Folk psychology re-assessed*, pp. 25–40. Dordrecht: Springer.