Cogito 1

Studies in Philosophy and its History

Cogito

Studies in Philosophy and its History

―――

Open to all philosophical genres and traditions, the book series *Cogito. Studies in Philosophy and its History* aims at publishing selected studies and texts with special focus on metaphysics, ethics, and their relations.
Cogito promotes a view of philosophy as an argumentative discipline, which needs to be both theoretically engaged and conscious of the historical dimension of theoretical discussions. Hence, the book series hosts readable and accessible texts, which contribute to philosophical advancement or which inquire into the historical roots of theoretical issues.
Cogito considers contributions from both senior and junior scholars, and promotes scientific dissemination by employing an open access policy. Submitted texts will be peer reviewed according to standard double-blind processes of evaluation. Contributions in English, Italian, French, Spanish and German language are eligible.

Questo volume è integralmente disponibile online a libero accesso nell'archivio digitale OpenstarTs, al link:
www.openstarts.units.it/dspace/handle/10077/14044

Practical Rationality in Political Contexts
Facing Diversity in Contemporary Multicultural Europe

edited by

Gabriele De Anna and Riccardo Martinelli

Table of contents

Rationality in Linguistic Interpretation: from Charity to Cooperativeness

PAOLO LABINAZ

1. Introduction

This paper deals with how human rationality is related to the practice of interpreting other people's linguistic behaviour. According to some well-known philosophers of language and mind, such as Willard V.O. Quine (1960), Donald Davidson (1984) and Daniel Dennett (1978, 1987), there is a necessary connection between our capacity to interpret other people's behaviour, be it linguistic or not, and our acknowledgement of their rationality. On these views, indeed, it would be very difficult (maybe impossible) to interpret other people's behaviour without also ascribing them a certain degree of rationality. Central to this claim is the Principle of Charity, according to which, roughly, maximizing agreement, or minimizing disagreement, with our interlocutors is a necessary condition of interpreting their linguistic behaviour successfully. That is, by assuming the speaker's rationality we can maximize agreement, or minimize disagreement, with her, making her linguistic behaviour understandable. However, if we put the Principle of Charity to work, some problems arise as to how this presumption of rationality is to be articulated. As I will try to argue, the Principle of Charity should be dismissed on the grounds that it makes difficult to determine what degree, if any, of rationality we have to attribute to our interlocutors in order to interpret their linguistic behaviour successfully.

A more promising way of accounting for the interpretation of other people's utterances may be found if we start from Paul Grice's Cooperative Principle, and his argumentative conception of rationality (Grice 1975/1989, 1991), as recently discussed by Marina Sbisà (2001, 2006). While according to the Principle of Charity, conformity to certain rational principles, whatever they are, must be ascribed *a priori* to our interlocutors, Grice's Cooperative Principle specifies what we should reasonably expect (*ceteris paribus*) from our interlocutors in conversation, insofar as we assume that they cooperate with us to reach a common goal, without imposing any specific rational principle to their linguistic behaviour. It is noteworthy that even when we believe that our interlocutors violate some of the expectations associated with the Cooperative Principle, we can still make sense of their linguistic behaviour insofar as we recognize that they have behaved in that way on the basis of some reasons, that is, in a rational way. On this view, rationality emerges from one's linguistic practices, particularly as a concern for the justification of one's own linguistic moves. I conclude by pointing out that studying rationality from the point of view of how we should interpret other people's speech sheds some light on how to conceive rationality itself. In particular, approaching rationality in this way gives us good grounds for moving from the received view that other people's rationality (as a conformity to norms) has to be assumed *a priori,* to an alternative view. This alternative view, which stands in need of further refinement, claims that rationality emerges in our concern to give reasons to support our behaviour, be it linguistic or of another nature.

2. Charity and rationality

Over the years, a wide debate has taken place concerning the nature of the Principle of Charity, and its application in linguistic interpretation. Roughly speaking, this principle states a general constraint on linguistic interpretation: to say that our interpretation of other people's linguistic behaviour must be charitable means that in order to interpret this linguistic behaviour successfully we must maximize agreement, or minimize disagreement, with them. As we will see, this requires in turn the assumption that the speaker being interpreted is, to a certain extent, rational, i.e. the interpreter ought to presume the speaker's rationality.

While the label "Principle of Charity" was coined by Neil L. Wilson (1959), the principle has been later reformulated and developed by several prominent philosophers, particularly Willard V.O. Quine and Donald Davidson.[1] Moreo-

---

[1]   Philosophers who have discussed (in different forms and for various reasons) the Principle of Charity in their philosophical theorizing are, among others, David Lewis (1974), Paul Thagard and Richard Nisbett (1983), Hilary Putnam (1975, 272-279).

ver, another well-known philosopher, Daniel Dennett, has developed a theory which, while differing in some respects from those of Davidson and Quine, holds that when predicting or explaining the behaviour of intentional systems we have to adopt a charitable attitude towards them, that is, we have to consider them as perfectly rational, or nearly so.

2.1 Quine on radical translation

In *Word and Object* (1960, 26-79), in order to argue for the indeterminacy of translation Quine elaborates a thought experiment in which a linguist is supposed to interact with a native speaker of an unknown community in an attempt to translate the latter's exotic language into her home language. Since the linguist has no prior knowledge of this language and there is no third shared language available, the only evidence available to the linguist is the native speaker's observable linguistic behaviour. Quine refers to this situation as a case of radical translation. Following Quine (1960, 29-30), suppose that when a rabbit appears the native speaker utters "gavagai." If this happens more than once, the linguist might conjecture that the expression "gavagai" refers to the rabbit, that is, can be translated into English as "gavagai" (or something like "Look, there is a rabbit"). However, in uttering "gavagai," the native speaker might refer to a portion of reality other than the rabbit considered as a whole (we suppose this is the portion of reality to which the English word "rabbit" actually refers), such as "rabbit parts," "temporal stage of a rabbit" or "instantiation of rabbithood," etc. Insofar as the linguist has only the observable linguistic behaviour of the native speaker at her disposal, she has no basis to decide what the most appropriate translation is. Indeed, considering one and the same linguistic stimulus, several hypotheses may be made. In addition, there is another question that the linguist has to tackle in translating the unknown language, that is, how to translate logical connectives (such as "∧", "∨" and so on) into the native speaker's language (Quine 1960, 57-61). For example, suppose that a native speaker assents to certain compound sentences that correspond to the logical form "p and not-p" (i.e., "the rabbit is on the grass and the rabbit is not on the grass"). Sentences with this form are absurd, or better contradictory, according to the semantic criteria of the linguist, as well as our own. But, as Quine points out, "[w]anton translation can make natives sound as queer as one pleases. Better translation imposes our logic upon them […]" (Quine 1960, 58). It is therefore suggested that the native speaker's utterance is not really contradictory and that the contradictory translation is a wrong translation.

In Quine's view, in the case of radical translation, since there may be an indefinite number of incompatible manuals of translation that are all consistent

with the same set of observable linguistic behaviour, the linguist must maximize her agreement with the native speaker. That is, only when the linguist assumes that there is a great deal of agreement between her and the native speaker, her task of translation can start. Agreement on the part of the linguist amounts to a charitable attitude towards the native speaker's observed linguistic behaviour. Quine proposes, among others, two methodological suggestions that have to be followed when translating sentences from an unknown language:

- "[…] fair translation preserves logical laws […]" (Quine 1960, 59);

- "[…] interlocutor's silliness, beyond a certain point, is less likely than bad translation – or, in the domestic case, linguistic divergence" (Quine 1960, 59), where silliness has to be evaluated from the linguist's perspective.

Put another way, Quine (1960, 58-59) suggests that the linguist has to normally consider an apparently nonsensical translation as the result of a mistake of the translator rather than irrationality or illogicality on the part of the native speaker. Charity requires that the linguist has to avoid translating the native speaker's utterances as contradictory or absurd. According to Quine, as seen previously, to avoid contradictory or absurd translation she has to "project" (in Quine's terms, "impose") her logic onto the native speaker.[2] However, this move can be understood in at least two different ways. While the first interpretation requires assuming that the native speaker's linguistic behaviour is logically consistent, regardless of whether she actually possesses a logical competence, the second one, which is a stronger assumption, requires the linguist to assume that the native speaker possesses a "real" logical competence including a certain amount of rules of first-order logic. Regardless of how one answers this question, which will be approached in Section 3, Quine claims that by adopting a charitable attitude towards the native speaker, the linguist should be able to carry over the truth from the sentences of the exotic language, be they occasion or compound, to her home language's sentences (Quine 1960, 60).

2.2 Davidson on radical interpretation

What Quine holds to be characteristic of genuine cases of radical translation, Davidson assumes that it also occurs in ordinary communication. According to him, what he calls "radical interpretation" is not limited to cases of translation but is concerned with interpreting utterances and ascribing propositional atti-

---

[2]   It is clear here that Quine (1960, 58) refers to a specific logical system, that is, first-order logic.

tudes in ordinary communication (Davidson 1973/1984, 128-129). On the basis of his analysis of radical interpretation, Davidson argues that in order to develop a theory of meaning, that is, a theory that determines the meaning or meanings of the sentences of a certain object language, we ought to establish a Tarski-style theory of truth for that language (Davidson 1973/1984, 130-131). According to Davidson, indeed, a theory of meaning based on Tarski's definition of truth "[…] provides an effective method for determining what every sentence means (i.e. gives the conditions under which it is true)" (Davidson 1966/1984, 8).

Radical interpretation differs fundamentally from radical translation in that the former is symmetrical while the latter is not: in radical interpretation, since both interlocutors want to understand each other, they share the same intention to communicate. However, as in the case of radical translation, in order to reach this goal both interlocutors have to maximize agreement, or minimize disagreement, between each other. Here Davidson refers to the Principle of Charity, but argues for a stronger form of it, according to which it is a necessary condition of the possibility of interpreting someone's speech and behaviour to assume that she is by and large rational and most of her beliefs are true (Davidson 1974a/1984, 183-198). While Quine uses the Principle of Charity as a methodological principle in translation, Davidson considers it as an *a priori* principle governing the ascription of propositional attitudes and the interpretation of linguistic behaviour (Davidson 1974/1984, 197). He claims that "if we are intelligently to attribute attitudes and beliefs, or usefully to describe motions as behaviour, then we are committed to finding in the pattern of behaviour, belief, and desire a large degree of rationality and consistency" (Davidson 1974b/1980, 237). So, when assigning meanings to a speaker's utterances, and contents to her attitudes, we must construe her as (i) being largely rational (or logical) and as (ii) having more true beliefs than false ones.[3] According to Davidson, there is a strong connection between (i) and (ii): insofar as most of our beliefs are true and there is a large degree of logical (or rational) consistency among them, new beliefs generated from those already available will be more likely true than false. If these two constraints are satisfied, we can identify a large degree of consistency and truthfulness in our interlocutors' utterances and thoughts, making them interpretable and so understandable.

---

[3]   In another work, Davidson adds a further interpretative constraint, which is not so relevant in the present context. This constraint requires the interpreter to assume that the speaker shares most of her own values, such as the desire "[…] to find warmth, love, security and success" (Davidson 1982/2004, 183).

2.3 Dennett on the intentional stance

According to Daniel Dennett, the assumption of rationality is a precondition of our ability to attribute intentionality to other beings (humans or not): "[t]he assumption that something is an intentional system is the assumption that it is rational" (Dennett 1971/1978, 10-11). In other words, according to Dennett, if something is taken to be an intentional system then she/he/it is must be assumed to be largely rational. Accordingly, in assuming that something thinks and acts rationally we adopt an "intentional stance" towards her/him/it (Dennett 1971/1978, 6-7). On the contrary, every time we depart from the assumption of rationality in trying to interpret other people's behaviour, we preclude ourselves from the possibility of explaining or predicting it. More specifically, when the intentional stance is adopted towards something

- an intentional attitude is first attributed to her/him/it,

- on the basis of the intentional attitude that has been attributed to her/him/it, it is possible then to predict her/his/its behaviour by determining what actions are rationally required in order that the intention be accomplished (Dennett 1971/1978, 6-9).

If she/he/it is assumed to be rational and consequently her/his/its propositional attitudes and their relationships to each other are taken to be so, we may be able to describe her/him/it in intentional terms. It is noteworthy that from Dennett's instrumentalist point of view the intentional stance is only a useful predictive tool, that is, relying on it helps us to predict what a rational individual will do given her goal(s) and cognitive situation, regardless of what actually happens in her mind. This means, in turn, that he understands intentional properties not to be real "mental" properties (see, e.g. Dennett 1971/1978, 16).

3. The presumption of rationality

The Principle of Charity has been a cornerstone in the debate concerning linguistic interpretation because, according to its supporters, "the position of the radical interpreter is the most fundamental position from which to investigate meaning and related matters, and it is needed to make sense of how the interpreter can see, on the basis of his evidence, another as a speaker" (Ludwig 2003, 17). However, there are still some relevant questions that need to be answered about the status of the Principle of Charity and its conditions of application. As to its status, the Principle of Charity can be interpreted in at least two different ways. On the one

hand, charity can be seen as a generalization of our best interpretative practices. On the other hand, it can be taken as a normative rule for interpretation. These two interpretations can be articulated in the following way respectively:

(Gen)   Successful interpretation of other people's linguistic behaviour is made according to the Principle of Charity.

(Nor)   Successful interpretation of other people's linguistic behaviour has to be made according to the Principle of Charity.

(Gen) is clearly a descriptive principle: it describes how successful interpretations take place, that is, what people actually do to interpret other people's linguistic behaviour successfully. On this view, the Principle of Charity is regarded as a heuristic device that allows us, most of the time, to interpret other people's linguistic behaviour successfully. While this interpretation of the principle may be attributed to a certain extent to Quine, according to Davidson the Principle of Charity plays a purely normative function, regardless of what people actually do in interpreting other people's linguistic behaviour. Indeed, since the correct interpretation always contrasts with many other interpretations that could have been made, there must exist "[…] an interpersonal standard of consistency and correspondence to the facts [that] applies to both the speaker and the speaker's interpreter, to their utterances and to their beliefs" (Davidson 1991/2001, 211). Davidson does not take a position on whether the Principle of Charity is actually valid in its particular applications, but considers it as a constitutive norm that governs any possible linguistic interpretation. On this view, (Nor) specifies what a successful interpretation of other people's linguistic behaviour amounts to. Since Davidson's interpretation of the Principle of Charity is central to the debate on linguistic interpretation in philosophy of language, in the remaining part of the paper we will refer to the Principe of Charity as (Nor). The next question, then, is: under what conditions does an interpreter comply with the Charity Principle in interpreting other people's linguistic behaviour?

As we have seen in the previous section, Quine, Davidson and Dennett agree that in interpreting other people's behaviour the most fundamental condition that must be respected is that of presuming their rationality. This condition is also known as "the presumption of rationality." According to some philosophers (and also some well-known neoclassical economists; see, e.g., Friedman 1953), the presumption of rationality is held to be so strongly entrenched in our ordinary life and not be, therefore, easily questionable (see, e.g., Rescher 1988, 191-194). In order to be able to understand your interlocutor, indeed, you have to assume that she will be on that occasion rational: no one would be able to understand a non-rational individual. As Donald Davidson (1991/2001, 211)

points out, "successful interpretation necessarily invests the person interpreted with basic rationality." That is, successful interpretations necessarily require that the speaker being interpreted is "imputed" with a certain degree of rationality. Since the presumption of rationality is an empty formula, we can make some suppositions as to what conditions it imposes on the interpreter (see also Thagard and Nisbett 1983, 251-252). Indeed, its injunction ("presume that…") may be interpreted according to different levels of stringency, such as the following:

(1) Presume that your interlocutors are sometimes rational.

(2) Presume that your interlocutors are usually rational.

(3) Presume that your interlocutors are always rational.

These requirements, which can be further decomposed into many others, show that the presumption of rationality can be applied with different levels of stringency. But what is the level of stringency adopted, either implicitly or explicitly, by Quine, Davidson and Dennett?

In the case of radical translation, the answer is not completely clear. On the one side, when Quine claims that "interlocutor's silliness, beyond a certain point, is less likely than bad translation," he seems to be committed to (2). By using the expression "beyond a certain point," Quine leaves open the possibility that in some specific circumstances we could prefer to assume the native speaker's illogicality or irrationality rather than trying to give sense to her utterances according to our logical criteria. On the other hand, in arguing for the projection of the linguist's logic onto the native speaker, Quine seems to assume a stronger presumption of rationality, which can be equated to (3). On this view, insofar we project our logic onto the native speaker, we can never translate her linguistic behaviour as contradictory, since she is assumed to be as logical (or rational) as we are. However, there is no real contradiction between these two positions: they can be taken to be two poles of a *continuum* with a number of intermediate positions that have to be determined contextually. When faced with a contradiction, supposing that the native speaker has assented to both a sentence and its negation, the linguist could either abandon her translation or decide that in consideration of the relevant evidence the native speaker's linguistic behaviour actually violates the principle of non-contradiction. If the latter case happens more than once, according to Quine (1960, 59) we are justified in concluding that the native speaker's language is untranslatable due to its non-logical structure. Since this type of case is very rare, as underlined by Rescorla (2013, 477), according to Quine it is enough for the linguist to not render the native speaker "[…] as denying obvious truths, whether those truths are logical, observational, or otherwise." However, obviousness is a very tricky concept that is usually as-

sociated with subjective or context-dependent standards. Therefore, in Quine's radical translation, the question of how stringent the presumption of rationality is remains underspecified.

In Davidson's theory of radical interpretation, the presumption of rationality is taken to be constitutive of the attribution of contents to utterances and thoughts. According to Davidson, indeed, we can attribute contents to the other person's utterances and thoughts only if we assume that they are interrelated in normatively appropriate ways, namely in rational ways. Therefore, the Principle of Charity, with its presumption of rationality, is not only a normative condition of interpretation, but also a constraint on any manifestation of rational behaviour. If we do not assume that a speaker's utterances and thoughts are interrelated in normatively appropriate ways, we cannot find rationality in her behaviour. On this view, the presumption of rationality supported by Davidson corresponds to (3), although, as he frequently underlines in his writings, local cases of irrationality are always possible (see, e.g., Davidson 1985/2004). What is fundamental for Davidson is that "[…] we cannot accept great or obvious deviations from rationality without threatening the intelligibility of our attributions" (Davidson 1980/2004, 156). This keeps open the question of what "great or obvious deviations from rationality" amount to and how such deviations are to be distinguished by minor or less obvious ones. In another passage, Davidson (1998/2005, 319) holds that "[c]harity is a matter of finding enough rationality in those we would understand to make sense of what they say and do, for unless we succeed in this, we cannot identify the contents of their words and thoughts." While Davidson continues to be vague about the degree of rationality required by charity (which is something to which we will come back in the next section), there seems to be also some sort of circularity in his reasoning. On the one hand, Davidson claims that insofar as we assume our interlocutors' rationality we can determine the contents of their utterances and thoughts. On the other hand, he argues that rationality can be attributed to people insofar as their utterances and thoughts are taken to be interrelated in normatively appropriate ways (Davidson 1980/2004, 156-157). But if we want to know whether utterances and thoughts are interrelated in normatively appropriate ways we have to know the contents of such utterances and thoughts, since the appropriateness of their relations depends also on their contents, not only on their linguistic or logical forms. Here Davidson seems to fall into a vicious cycle: linguistic interpretation requires attribution of rationality to the speaker being interpreted, but, at the same time, attribution of rationality presupposes some level of interpretation of the speaker's utterances and thoughts, since one needs to know whether they are interrelated in normatively appropriate ways, that is, whether they are rational. Can this vicious circle be broken? Davidson' general tendency seems to be that of considering the

presumption of rationality as the most basic precept we have in linguistic interpretation. In this sense, on his account, without attribution of rationality, we have no linguistic interpretation.

Dennett's way of conceiving the presumption of rationality is similar in many respects to that of Davidson. In this sense, they face similar challenges with respect to the presumption of rationality. As Dennett (1978, 11) points out, "[…] one gets nowhere with the assumption that entity x has beliefs p, q, r, … unless one also supposes that x believes what follow from p, q, r, …; otherwise there is no way of ruling out the prediction that x will, in the face of its beliefs p, q, r, … do something utterly stupid […]." Here Dennett presupposes level (3) of stringency: assuming people's rationality underpins the possibility of explaining and predicting their intentional behaviour.

According to Davidson and Dennett, and to a certain interpretation of Quine's analysis, the Principle of Charity requires that the presumption of rationality must be applied with a strong level of stringency, that is, according to the level (3). On this view, to presume other people's rationality, conceived as a conformity to certain norms (whatever they may be), is a necessary condition of the possibility of interpreting their linguistic behaviour successfully.

4. Degrees and norms of rationality

In the previous section, we have assumed, following Davidson, that the Principle of Charity requires a normative reading and that such a normativity is constitutive of any linguistic interpretation. Moreover, we have argued that to presume the speaker's rationality is a necessary condition of interpreting her linguistic behaviour successfully. However, some clarifications are still needed as regards the normative function of the Principle of Charity and the validity of its application in specific circumstances. In particular, two main questions can be posed:

(i) what degree of rationality do we have to attribute to a speaker in order to make her linguistic utterances understandable?

(ii) what are the norms of rationality compliance which have to be attributed to the speaker?

With respect to (i), as underlined in the previous section, the question arises due to the lack of clarity in Davison's claims (but also in those of Quine and Dennett) about how the degree of rationality to be attributed to our interlocutors should be determined: Davidson speaks of "a large degree of rationality," "basic rationality" and "enough rationality" etc. How could an appropriate degree of rationality

be determined in the speaker in specific circumstances? This problem does not seem to affect the validity of Davidson's theory in any way: indeed, since according to him the Principle of Charity can be taken as a necessary constraint on all possible interpretation, its general validity cannot be questioned by its particular application in a given context (Davidson 1974/1984, 197). That is, one question is the general validity of Davidson's account, while another is its applicability in specific cases. Davidson claims to be interested in approaching only the former question, considering the latter to be irrelevant in his attempt to develop a general theory of meaning (see Section 2.1). Therefore, since the Principle of Charity is a general constraint on linguistic interpretation, we should not care whether it may be applied to specific situations or not. However, even if this can be accepted for the Principle of Charity, the same reasoning cannot be applied to the presumption of rationality which the Principle of Charity requires from interpreters. Since this presumption must have a specific content, it cannot be as general and abstract as the Principle of Charity is supposed to be. Indeed, any individual can be said to be more or less rational: people's rationality can be seen as a matter of degree. But which is the minimum degree of rationality to be attributed to the speaker in order to perform a successful interpretation of her linguistic utterance? This question leads us to (ii), the issue of what norms of rationality we should assume our interlocutors conform to.

As to (ii), the question is how to identify the appropriate norms of rationality that are to be imputed to our interlocutors (see also Lukes 1982, 263). Since it is a platitude that humans are not perfectly rational in ordinary life, one could ask what norms of rationality are actually supposed to guide our behaviour. Every day we experience our failures, as well as those of others, in reasoning activities because of prejudices, laziness, carelessness, lack of time, inattention etc. While failures of this kind are due to momentary lapses, there are many others that are deeply rooted in our cognition, and of which we are not always aware. As many experimental studies on reasoning and decision-making suggest, we frequently tend to deviate from classical norms of rationality due to our reliance on so-called heuristics and biases (see, e.g., Gilcovich *et al.* 2002). But how could the Principle of Charity and its presumption of rationality be led to cope with these facts that are apparent to anyone in ordinary situations, as well as in experimental labs? One could argue that it is not only wrong to present the Principle of Charity as a generalization of our interpretative practices, but also, as Davidson does, to justify it in normative terms, that is, as a normative constraint on interpretation. Davidson can reply to this objection by arguing that although people are locally irrational, they can be judged so on the assumption that they are usually rational: "irrationality is a failure within the house of reason" (Davidson 1982/2004, 169). As observed by Rescorla (2013, 478), according to Davidson examples of irra-

tionality can be seen as "[…] *deviations* from background conformity to rational norms." But what are such norms? In a reply to some commentators, Davidson (1998/2005, 319) claims that "[t]hese norms include norms of logical consistency, of action in reasonable accord with essential or basic interests, and the acceptance of views that are sensible in the light of evidence." If the Principle of Charity requires the presumption that our interlocutors have to be taken as having logically consistent beliefs and probably consistent probability assignments, behaving in order to maximize their basic interests, taking into account all relevant information in arriving at conclusions, many of Davidson's critics, such as Cristopher Cherniak, Alvin Goldman and Stephen Stich, are ready to reply that these requirements present a number of problems, such as the following:

- If the results of reasoning and decision making experiments are correct, it is not completely clear whether humans may be assumed to conform to these requirements as closely as Davidson claims.[4] As a matter of fact, people often make decisions by which they gain no benefit, contradicting the assumption that they behave in order to maximize their basic interests, and utter contradictory claims or fail to infer conclusions, which can be seen as a sign of their illogicality. According to Cherniak (1986), the kind of rationality that is applicable to creatures such as humans is a minimal rationality. Therefore, requirements, such as those posed by Davidson, cannot really be considered "normative" for rational beings as we are, since we are not able to respect them. In this sense, attributing this kind of "ideal" rationality to our interlocutors is not a good move in linguistic interpretation. In one of his examples, Cherniak (1986) refers to issues of checking one's beliefs for consistency, which is one of the requirements suggested by Davidson. According to Cherniak, a complete check of one's current set of beliefs for truth functional consistency would lead to a computational explosion.[5] It is usually the case, indeed, that when an individual tries to solve contradictions between her beliefs she does not commit herself to a process of exhaustive search. People are faced with temporary inconsistency continuously, whenever something unexpected occurs. Accordingly, one of the main requirements posed by Davidson, that is, logical consistency among one's beliefs, seems to be too high a standard of rationality to be safely attributed to any speaker. More generally, it is not

---

4   It is noteworthy that according to Dennett (1973/1978) when irrationality takes place (but according to which standard?) no prediction of the other's behaviour can be made by adopting the intentional stance.

5   Cherniak (1986, 93-94) has calculated that, in order to make a complete check for truth functional consistency a belief system containing 138 logically independent propositions would require the power of a super computer which would need to work (in order to complete this task) for more than twenty billion years.

clear how to distinguish between norms of rationality that are appropriate for the kind of finite beings we are and those that are beyond human cognitive capacity.

- Another relevant issue concerns the question of whether rationality and its norms are universal, that is, whether they apply to all humans, regardless of their belonging to a specific community or society. Some recent experimental studies suggest that people from different cultures exhibit different ways of thinking, and different styles of reasoning in particular, suggesting that people's cognitive systems may differ significantly in some basic cognitive processes, such as categorization, perception and reasoning (for a survey of these studies, see Nisbett 2003). What should an interpreter do when interacting with people who have different basic cognitive processes, particularly those underlying reasoning? Does she have to follow Quine's suggestion and thus project her "logic" onto her "exotic" interlocutors? A good answer could be that the interpreter has to project onto her "exotic" interlocutors only a part of her logic, that is, the part that is common to all humans. However, although there must probably exist "a 'bridgehead' of true and rational beliefs" (Hollis 1982, 73) that are common to all humans, it appears that we are far from identifying it.

A possible solution to the issues just mentioned is to replace the Principle of Charity with some other principle that is not grounded on the presumption of rationality. As Richard Grandy (1973) observed, since there can be successful interpretations even when cases of irrationality take place, whether they are real or apparent, we can hypothesize that the presumption of rationality does not ground our mutual understanding. Consequently, Grandy argues for what he calls the "Principle of Humanity." While the Principle of Charity requires the presumption of rationality, Grandy's principle requires the application of a process of "simulation." According to the Principle of Humanity, when interpreting other people's behaviour one has to suppose that one has the same beliefs and desires as those of the speaker in order to ask oneself (*as if* one was her) what she would do in that case. What matters in this case is that other people share with us the same "[…] pattern of relations among beliefs, desires and the world" (Grandy 1973, 443). However, since Grandy remains vague as to what such a pattern of relations amount to, the interpreter could in principle impose any structure upon other people's beliefs and desires, regardless of whether this structure is subjective, context-dependent, logical and so on.

In sum, the centrality of the presumption of rationality in linguistic interpretation seems to be overestimated and, at the same time, its formulation too vague to be acceptable, regardless of whether we should take the Principle of Charity as normative or not. Moreover, when the presumption of rationality is articulated,

as in the case of Davidson, it is very hard to find a definite set of norms of rationality that can be imputed to our interlocutors. Therefore, insofar as the Principle of Charity requires the presumption of rationality, it cannot be accepted as a necessary condition of linguistic interpretation.

5. Linguistic interpretation, Grice's Cooperative Principle and argumentative rationality

As we have shown in the previous section, there are serious doubts that rationality is the key to interpreting other people's linguistic behaviour. It may happen that some linguistic behaviour deviates from the relevant norms of rationality, while others meet such norms in some contexts but not in others, and this variability in people's linguistic behaviour compliance with norms of rationality does not help us to maximize agreement, or minimize disagreement, with them.

Accordingly, we should investigate whether there are other, more plausible, ways of making sense of other people's linguistic behaviour. How can we maximize agreement, or minimize disagreement, with our interlocutors in interpreting their linguistic behaviour, while leaving aside the presumption of rationality?

Obviously there could be more than one way of approaching this question, some of which are based on normative considerations (as in the case of Davidson and Dennett), while others are based on psychological ones (see, e.g., Stich 1984; Goldman 1989). Here I will try to outline an alternative approach that aims to account for the relation between linguistic interpretation and rationality, which is based on Paul Grice's Cooperative Principle and his argumentative conception of rationality (see Grice 1975/1989, 1991). At first sight the Principle of Charity and Grice's Cooperative Principle seem to differ in many respects,[6] in particular, it has been argued that the latter cannot be as fundamental in linguistic interpretation as the former is taken to be. Nevertheless, we will try to show that there are good reasons for taking Grice's Cooperative Principle as "[…] a presumption that receivers make (and that it is rational for them to make) for the sake of giving as full an interpretation as possible of what they are told" (Sbisà 2006, 234). As suggested by Sbisà (2001, 2006), this way of conceiving the Cooperative Principle requires an analysis of Grice's notions of non-natural meaning and "implicature" from the hearer's perspective. Within this framework, I maintain that in order to interpret other people's linguistic behaviour we do not need to conceive of their rationality as a conformity to certain norms. Rather, we have to rely upon a cooperativeness presumption regarding the speaker's linguistic

---

[6]   For a comparison between Grice's and Davidson's philosophy of language, see Avramides (2001).

behaviour in conversation. Since according to Grice (1975/1989, 28) talking can be seen "[…] as a special case or variety of purposive, indeed rational, behaviour," rationality can be found in our linguistic practices, insofar as these embody a concern for the justification of one's own linguistic moves.

5.1 Grice's Cooperative Principle and linguistic interpretation

Let me begin with Grice's formulation of the Cooperative Principle: "[m]ake your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" (Grice 1975/1989, 26). According to the standard interpretation, the Cooperative Principle, together with the conversational maxims that specify its content, is taken to suggest what speakers should do, and probably sometimes do, in order to be cooperative in conversation. This "normative" reading rests on Grice's formulation of the principle that includes the verb "to make" in the imperative mood (second person) and so can be equated to an order: "you ought to do so and so." However, as observed by Sbisà (2006, 234), Grice claims that in "calculating" a conversational implicature what matters is "[…] the assumption, or presumption, that the speaker is observing the Cooperative Principle and not the real observance of it, nor the (possible) obligation, or commitment, of the speaker to observe it." In other words, if this reading is correct, Grice does not treat the Cooperative Principle as a norm that has to be followed by every speaker when participating to a conversation, but as a presumption the audience makes in order to understand what has been said. As Grice (1975/1989, 30) points out,

> […] observance of the Cooperative Principle and maxims is reasonable (rational) along the following lines: that anyone who cares about the goals that are central to conversation/communication […] must be expected to have an interest, given suitable circumstances, in participation in talk exchanges that will be profitable only on the assumption that they are conducted in general accordance with the Cooperative Principle and the maxims.

Here we can find two interesting suggestions made by Grice that may help to overcome the Principle of Charity and its presumption of rationality:

(i) In order to maximize agreement, or minimize disagreement, with other speakers we must expect that they, or at least those who are presumed to care about the goals of the conversation, have some interests in taking part in it and so in communicating with us.

(ii) When (i) holds, as long as we are interested in making the conversation profitable, we should assume that such a conversation is conducted in accord-

ance with the Cooperative Principle and its conversational maxims, that is, we must presume that every participant, including ourselves, will behave in accordance with them (or at least with the Principle governing them).

As Grice points out in the quotation above, the expectation specified in (i), while being necessary, may have some value only if the presumption suggested in (ii) takes place: a conversation is profitable as long as all the participants take each other as being cooperative. In order to have mutual understanding in a conversation, it is not enough to ascribe to the other participants a concern for communicating with us, what matters is to presume that their linguistic moves are made in accordance with the Cooperative Principle and its conversational maxims. In doing so, we acknowledge them as rational beings, since their linguistic moves are taken to be made on the basis of some reasons, that is, they are taken to be appropriate to achieve "the accepted purpose or direction of the talk exchange" (Grice 1975/1989, 26). Accordingly, whether our interlocutors observe the Cooperative Principle is not strictly an empirical question but a reasonable decision one makes: thanks to this decision, indeed, we make ourselves able to understand their utterances in the ongoing conversation.

By drawing a parallel with the Principle of Charity and its presumption of rationality, we might go on to consider the attribution of the observance of the Cooperative Principle to the speaker (involving the attribution of the actual observance or open violation of the conversational maxims). As underlined by Sbisà (2001, 198) the Cooperative Principle "[…] should be considered as a complex assumption that the audience makes about an utterer, in order to fully understand what the utterer means$_{nn}$ [non-naturally]." According to Grice (1975/1989, 26), indeed, by assuming that the speaker is cooperative, we can determine what we should expect (*ceteris paribus*) from her utterances. What matters here is not the intention of the speaker, which is standardly considered to determine the presence of non-natural meaning and of implicatures, but the ascription of intentions we make in correspondence to the meaning we take to be conveyed by her utterances. Obviously, as underlined by the use of the expression "ceteris paribus," the assumption that our interlocutors are cooperative can be cancelled if we believe that there are sufficient reasons for doing so. When (if ever) this cancellation is complete, linguistic communication fails.

Something different happens when, provided that the speaker is taken to be cooperative, we believe that her utterance violates one or more conversational maxims. When some conversational maxim is violated, but we assume that speaker is cooperative, a conversational implicature has to be worked out in order to achieve the correct understanding of what the speaker means. In such a case, one can ask why the speaker violates one or another maxim if she intends

to be cooperative. Insofar as we can recognize that the "cooperative" speaker has some reasons for violating a conversational maxim, we can ascribe to her the intention that in violating such a maxim she has meant to convey something more than what she has literally said. As Grice points out (1975/1989, 39-40), indeed, "to calculate a conversational implicature is to calculate what has to be supposed in order to preserve the supposition that the Cooperative Principle is being observed."

Successful interpretations of other people's utterances is then determined by our expectation that such utterances have been made in accordance with the Cooperative Principle and its conversational maxims. Since, according to Grice, the meaning conveyed by linguistic expressions can be worked out when participants in the conversation take each other as cooperative, the Cooperative Principle can be seen as setting out a general constraint on linguistic interpretation.

5.2 From linguistic interpretation to argumentative rationality, and back again

According to the supporters of the Principle of Charity, at least in its most popular version (see Section 2), the speakers' rationality, conceived as a conformity to certain norms, must be presupposed in order to interpret their linguistic behaviour correctly. However, as we have seen in Section 4, relying on the Principle of Charity in linguistic interpretation gives rise to certain troubles due to the difficulties in defining the extent and degree of rationality needed and the norms which have to be imputed to the speaker. In contrast with this view, according to the Gricean approach here proposed, people's rationality emerges in their linguistic practices, as a concern for their linguistic moves as justified. This alternative view on the relation between linguistic interpretation and rationality is based on a conception of rationality, which Sbisà (2006; 2007) has labelled "argumentative," that can be found in Grice's Carus Lectures (1983), published posthumously with the title *The Conception of Value* (1991). According to this conception, as Sbisà rephrases it (2006, 241-242): human rationality consists in "[...] a concern that one's moves are justified and a capacity (to some degree) to give effect to that concern" (see also Grice 1991, 82-83). The argumentative conception of rationality comprises two aspects: while the first aspect is concerned with one's motivations for justifying what one says or does, the second one focuses on one's actual ability to give effect to this concern, that is, to provide reasons supporting one's behaviour, be it linguistic or not. Accordingly, it implies that it is not enough to be able to provide reasons supporting one's moves, but one must also care about actually having such reasons, that is, one has to find it valuable to provide reasons supporting one's claims or actions.

Let us consider why this conception of rationality matters in linguistic interpretation. Insofar as we take other participants in the conversation as behaving in accordance with the Cooperative Principle, we consider them as being motivated to communicate with us, that is, their linguistic moves are not made at random but based on reasons, i.e. rationally justified, even when one or more conversational maxims are violated. In this sense, both Grice's Cooperative Principle and his conception of argumentative rationality converge on the idea "[t]he justification of one's moves remains a task for every human being who wants to take him or herself seriously as a person" (Sbisà 2006, 243). It is clear here that human rationality is not to be described as a conformity to certain norms, but is something that one can find in conversation, if considered as a cooperative enterprise. Given this general outlook, we turn now to the role of argumentative rationality in linguistic interpretation. On the one side, in presuming that the speaker observes the Cooperative Principle and its conversational maxims, the hearer considers her as behaving rationally, that is, that her linguistic behaviour is well grounded, and appropriate in trying to reach "the accepted purpose or direction of the talk exchange" (Grice 1975/1989, 26). On the other side, when some of the conversational maxims are taken to be openly violated, in the attempt to give sense to the speaker's utterances the hearer may still assume that she could be taken to be cooperative, after all. When this happens, the hearer can grasp what the speaker has meant by assuming that such violations have been made on the basis of some reasons. As observed by Sbisà (2006, 243-244), indeed, "[e]xplaining a certain part of the meaning of an utterance as a conversational implicature [...] requires that there be an inferential path leading to that implicature and thus an argument in support of it."

More generally, ascription of cooperativeness is a fundamental aspect of our sociality because even when our interlocutors' utterances do not satisfy some of the expectations made explicit by the conversational maxims – e.g. what has been said is taken to be false or lacking adequate evidence, is not as informative as is required, is not relevant or perspicuous – insofar as we acknowledge them as cooperative, we are still able to work out the global meaning conveyed by their utterances. On the one side, in taking our interlocutors as cooperative, we acknowledge them as persons who have reasons for what they say and do, that is, as persons endowed with rationality. On the other side, as we have seen before, since the Cooperative Principle can be considered as a general condition of linguistic interpretation: in doing so we make our interlocutor's utterances interpretable and so understandable.

6. Concluding remarks

As stated at the beginning of the paper, my main concern here was to discuss the relation between linguistic interpretation and rationality. In the first part, I examined some relevant accounts of linguistic interpretation based on charity, particularly those proposed by Quine, Davidson and Dennett, respectively. Each of these assumes that in order to interpret other people's linguistic behaviour successfully we have to assume *a priori* their rationality, conceived as a conformity to certain norms. As we have seen, however, it is very hard, or maybe impossible, to determine the extent and degree of rationality needed and the norms which have to be imputed to the speaker. In contrast, we have seen that insofar as we take cooperation as a precondition for linguistic interpretation, we can make sense of other people's utterances by assuming that they are made on the basis of reasons, that is, in a rational way. While this argumentative conception of rationality requires further clarification and refinement, as regards why we should have concern for justifying our linguistic moves, what our ability to give effect to this concern amounts to, what a good reason is etc., relying on it helps us to consider human rationality from a different perspective. In particular, given the strict relationship between the attribution of cooperativeness to a speaker and the acknowledgement of her rationality in conversation, people's rationality can be said to emerge in their linguistic practices. Indeed, insofar as linguistic interpretation requires attribution of cooperativeness to be performed successfully, in presuming that our interlocutors' utterances have been made in accordance with the Cooperative Principle, we treat them as having a concern for justifying their linguistic moves, that is, as having concern for acting rationally in conversation.

References

Avramides, A. 2001. "Davidson, Grice, and the Social Aspects of Language." In *Paul Grice's Heritage*, edited by Giovanna Cosenza, 115-138. Turnhout: Brepols.

Cherniak, Christopher. 1986. *Minimal Rationality*. Cambridge, MA: The MIT Press.

Davidson, Donald. 1966. "Theories of Meaning and Learnable Languages." In *Logic, Methodology, and Philosophy of Science: Proceedings of the 1964 International Congress*, edited by Yehshoua Bar-Hillel, 383-394. Amsterdam: North Holland. Reprinted in D. Davidson (1984), *Inquiries into Truth and Interpretation*, 3-16. Oxford: Oxford University Press.

Davidson, Donald. 1973. "Radical Interpretation." *Dialectica* 27: 314-328. Reprinted in D. Davidson (1984), *Inquiries into Truth and Interpretation*, 25-140. Oxford: Oxford University Press.

Davidson, Donald. 1974a. "On the Very Idea of a Conceptual Scheme." In *Proceedings and Addresses of the American Philosophical Association 47:* 5-20. Reprinted in D. Davidson (1984), *Inquiries into Truth and Interpretation*, 183-198. Oxford: Oxford University Press.

Davidson, Donald. 1974b. "Psychology as Philosophy." In *Philosophy of Psychology,* edited by Stuart Brown, 41-52. New York: Barnes & Noble. Reprinted in D. Davidson (1980), *Essays on Actions and Events*, 229-239. Oxford: Oxford University Press.

Davidson, Donald. 1980. "Toward a Unified Theory of Meaning and Action." *Grazer Philosophische Studien* 11: 1-12. Reprinted as "Unified Theory of Thought, Meaning, and Action." In D. Davidson (2004), *Problems of Rationality*, 151-166. Oxford: Clarendon Press.

Davidson, Donald. 1982. "Paradoxes of Irrationality." In *Philosophical Essays on Freud*, edited by Richard Wollheim and James Hopkins, 289-305. Cambridge: Cambridge University Press. Reprinted in D. Davidson (2004), *Problems of Rationality*, 169-187. Oxford: Clarendon Press.

Davidson, Donald. 1985. "Incoherence and irrationality." *Dialectica* 39: 345-354. Reprinted in D. Davidson (2004), *Problems of Rationality*, 189-198. Oxford: Clarendon Press.

Davidson, Donald. 1991. "Three varieties of knowledge." *Royal Institute of Philosophy Supplement* 30: 153-166. Reprinted in D. Davidson (2001), *Subjective, Intersubjective, Objective*, 205-220. Oxford: Oxford University Press.

Davidson, Donald. 1998. "Replies." *Crìtica* 30. Reprinted as "Appendix: Replies to Rorty, Stroud, McDowell, and Pereda (1998)," in D. Davidson (2005), *Truth, Language and History: Philosophical Essays*, 315-327. Oxford: Clarendon Press.

Dennett, Daniel. 1971. "Intentional Systems." *Journal of Philosophy* 78: 87-106. Reprinted in D. Dennett (1978), *Brainstorms: Philosophical Essays on Mind and Psychology*, 3-22. Montgomery, VT: Bradford Book.

Dennett, Daniel. 1973. "Mechanism and Responsibility." In *Essays on Freedom of Action*, edited by Ted Honderich, 157-184. London: Routledge. Reprinted in D. Dennett (1978), *Brainstorms: Philosophical Essays on Mind and Psychology*, 233-255. Montgomery, VT: Bradford Book.

Dennett, Daniel. 1987. *The Intentional Stance*. Cambridge, MA: The MIT Press.

Friedman, Michael.1953. *Essays in Positive Economics*. Chicago: University of Chicago Press.

Gilovich, Thomas, Dale Griffin, and Daniel Kahneman, eds. 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment*. New York: Cambridge University Press.

Goldman, Alvin. 1989. "Interpretation Psychologized." *Mind and Language* 4: 161-185

Grandy, Richard. 1973. "Reference, Meaning and Belief." *Journal of Philosophy* 1973: 439-452.

Grice, Paul H. 1975. "Logic and Conversation." In *Syntax and Semantics 3: Speech Acts*, edited by Peter Cole and Jerry Morgan, 41-58. New York: Academic Press. Reprinted in P. Grice (1989), *Studies in the Ways of Words*, 22-40. Cambridge, MA: Harvard University Press.

Grice, Paul H. 1991. *The Conception of Value*. Oxford: Oxford University Press.

Hollis, Martin. 1982. "The Social Destruction of Reality." In *Rationality and Relativism*, edited by Martin Hollis and Steven Lukes, 67-86. Cambridge, MA: The MIT Press.

Lewis, David. 1974. "Radical Interpretation." *Synthese* 23: 331-344.

Ludwig, Kirk. 2004. "Rationality, Language, and the Principle of Charity." In *The Oxford Handbook of Rationality*, edited by Alfred Mele and Piers Rawling, 343-362. New York: Oxford University Press.

Lukes, Steven. 1982. "Relativism in its Place." In *Rationality and Relativism*, edited by Martin Hollis and Steven Lukes, 261-305. Cambridge, MA: The MIT Press.

Nisbett, Richard. 2003. *The Geography of Thought: Why We Think the Way We Do*. New York: The Free Press.

Putnam, Hilary. 1975. "Language and reality." In *Mind, Language and Reality: Philosophical Papers, vol. 2,* 272-279. Cambridge, MA: Cambridge University Press.

Quine, Willard V.O. 1960. *Word and Object*. Cambridge, MA: The MIT Press.

Rescorla, Michael. 2013. "Rationality as a Constitutive Ideal." In *A Companion to Donald Davidson*, edited by Ernie Lepore and Kirk Ludwig, 472-488. Oxford, UK: John Wiley & Sons.

Rescher, Nicholas. 1988. *Rationality*. New York: Oxford University Press.

Sbisà, Marina. 2001. "Intentions from the Other Side." In *Paul Grice's Heritage*, edited by Giovanna Cosenza, 185-206. Turnhout: Brepols.

Sbisà, Marina. 2006. "Two Conceptions of Rationality in Grice's Theory of Implicature." In *Rationality of Belief and Action*, edited by Elvio Baccarini and Snježana Prijić-Samaržija, 233-247. Rijeka: University of Rijeka.

Sbisà, Marina. 2007. "On Argumentative Rationality." *Anthropology & Philosophy* 8: 89-99.

Stich, Stephen. 1984. "Relativism, Rationality and the Limits of Intentional Description." *Pacific Philosophical Quarterly* 65: 211-35.

Thagard, Paul and Richard Nisbett. 1983. "Rationality and Charity." *Philosophy of Science* 42: 379-394.

Wilson, Neil L. 1959. "Substances without Substrata." *The Review of Metaphysics* 12: 521-539.