RATIONALITY, LANGUAGE, AND THE PRINCIPLE OF CHARITY

KIRK LUDWIG

Making sense of the utterances and behaviour of others, even their most aberrant behaviour, requires us to find a great deal of reason and truth in them.

—Davidson 1984, chap. 10, 153

1. INTRODUCTION

THIS chapter deals with the relations between language, thought, and rationality, and especially the role and status of assumptions about rationality in interpreting another's speech and assigning contents to her psychological attitudes—her beliefs, desires, intentions, and so on. Central to the discussion below will be the

status, in particular, of the Principle of Charity, first introduced by W. V. Quine as a maxim of translation, "assertions start[l]ingly false on the face of them are likely to turn on hidden differences of language" (Quine 1960a, 58–60). Donald Davidson has advocated a stronger form of the principle, which enjoins as necessary for interpretation of another's speech the assumption that she is largely rational and has largely true beliefs (Davidson 1984, 27, 136–37, 152–53, 159, 168–67, 196–97, 200–201).

The discussion will be organized around the following three questions:

What is the relation between rationality and thought?
What is the relation between rationality and language?
What is the relation between thought and language?

These questions are not independent. To possess a language is to be able to speak to another and to understand another's speech. One must therefore be an agent, something capable of acting, as opposed to merely moving or being moved, to possess a language. Language therefore presupposes thought and action. If rationality is a condition on thought and agency, as is widely (though not universally) assumed, then it is likewise a condition on possessing a language. In this case, seeing another as a potential interlocutor carries a commitment to finding her to be fundamentally a rational being, and to regarding oneself as likewise fundamentally a rational being. On the other hand, there is a long tradition that sees language as essential for rationality, but not for thought generally. Aristotle famously defined man as "the rational animal" (1984: *Topics*, book 5, 132a22–132a27; *Nicomachean Ethics* 1.7.1097b22–1098a20; *De Anima*, 1.5.645b14). We occupy an even more privileged position if, as Davidson has controversially argued (Davidson 1984, chap. 11; 2001, chap. 7), language is a condition on thought, and rationality is essential to both: then thought, rationality, and language are possessed altogether or not at all.

Section 2 takes up the relation of rationality to thought. Section 3 discusses the relation between rationality and the power of speech. Section 4 takes up the relation of thought to language. Section 5 summarizes the discussion.

2. RATIONALITY AND THOUGHT

It is widely accepted that rationality is essential for thought. This section explains what this view comes to, the reasons for it in outline, and some of the objections that have been advanced against it. We begin with a brief characterization of what

is meant by "thought" in this discussion, and then of the domain and require-
ments of rationality.

The term "thought" will be used to cover any psychological attitude with a
propositional content. The term "propositional attitude" (coined by Bertrand Rus-
sell ([1922] 1961; [1918] 1985) will also be used interchangeably with "thought" in
this sense. Central examples are beliefs and desires. *A*'s belief that he is handsome
has as its content that he is handsome; his desire to be admired has as its content
that he is admired. Propositional attitudes are individuated by their psychological
modes and contents. Thus, different attitudes can have the same content if they
are entertained in different modes: one may have a belief that one will get well,
for example, as well as a desire to get well. Other examples of propositional
attitudes are intending, hoping, fearing, considering, wishing, and doubting. These
are the psychological states especially relevant to a discussion of rationality because
their contents, being propositional, can bear logical and semantic relations to one
another; for example, one propositional content can require or support, or be
incompatible or inconsistent with, the truth of another.

From antiquity, the domain of rationality has been divided into the theoret-
ical, having to do with the formation of belief, and the practical, having to do
with the expression of agency (the terminology is due to Aristotle; for contem-
porary discussions see Audi 2001 and Harman 1999, chap. 1). Theoretical ration-
ality aims at arriving at true belief and avoiding false belief, nonhaphazardly.
Practical rationality, to which theoretical rationality is an important aid, aims at
getting what one most wants, in accordance with one's beliefs about what one
can get and how one can get it (and, perhaps, though controversially, with eval-
uating one's ultimate ends; see, e.g., Audi 1990a and Brandt 1998). The degree to
which someone is rational depends on the degree to which his attitudes exhibit
patterns at and across times appropriate for ideal pursuit of his theoretical and
practical goals.

Theoretical rationality is concerned with having representational states that
exhibit coherence at a time in the sense particularly of not displaying patterns
that frustrate the goal of having true beliefs and avoiding false ones. Thus, for
example, consistency in what one believes is an obvious goal of full or ideal
rationality. Recognized inconsistency is worse than unrecognized inconsistency,
though, and in cases in which it is difficult to discover the inconsistency we do
not ordinarily count someone as irrational. (Frege's failure to recognize the in-
consistency of his axioms for arithmetic does not convict him of irrationality.)
Similarly for holding beliefs that, in the light of one's evidence, are not likely to
be jointly true. Theoretical rationality concerns also how new beliefs are acquired
in the light of new evidence, and with reasoning from, or acquiring new beliefs
in the light of, beliefs which one already has. In general, the goal is to acquire
true beliefs about, or relevant to, what one is interested in and to avoid false
beliefs. Often what rationality requires is thrown into clearer relief by its break-

downs. Thus, wishful thinking, believing something because you want it to be true, and arbitrary belief formation, believing for no good reason, are irrational, while apportioning belief to the degree of evidence is rational. (See chap. 1, this volume.)

Practical rationality is also concerned with patterns of attitudes relevant to action, centrally belief, desire, and intention, both at and across times. An intransitive preference ranking is an example of an irrational pattern among conative states, since it can lead systematically to the frustration of one's practical interests (see chap. 10, this volume). If you prefer A to B, B to C, but C to A, then in principle you can be led to trade something of value (a penny, for example) in an endless cycle to get B for C, A for B, C for A and then again B for C, and so on, each pairwise trade seeming rational, though the entire set is not. Practical rationality concerns also the effective coordination of desires and beliefs in the pursuit of one's ends, which requires that one recognize what are the best means to ends most preferred and then implement them. Doing what one does not judge best all things considered—weakness of the will—is a familiar breakdown of diachronic practical rationality (see chap. 13, this volume, and Davidson 1980, chap. 2). Similarly, though there is no general requirement on consistency in what one desires, there is a requirement on consistency in what one intends or plans to do, since inconsistent plans (the result of desires put through the sieve of practical reasoning) cannot be conjointly carried out (Bratman 1987, chap. 8).

Having the power of thought and action obviously does not require perfect rationality, whatever that could come to. Most of us are subject to all too familiar failings in both reasoning and acting. This gives point to seeing rationality as a normative requirement, as a standard by which to judge our thought and behavior. The question whether rationality is required for thought is whether something can be a thinking being without being *largely* rational, or, more generally, what the limits are on how irrational one can be and still be seen as capable of thought. Thus, the thesis that thought requires rationality can be put as the thesis that propositional attitudes can appear only in largely rational patterns, synchronic and diachronic. This is to say that the normative requirements of rationality, which tell us how we ought to reason, deliberate, and act, are also descriptive requirements on what it is to be a thinking being: we think and act largely as we ought, or we do not do so at all.

The case for rationality being a requirement on thought rests on reflection on the conditions under which it is appropriate to attribute to something the basic attitudes of belief and desire, which are the primitive ingredients of agency. (We will assume that beliefs and desires come together or not at all—that is, that all thinkers are agents. While this might be challenged, there will not be space to discuss it adequately here.)

Beliefs come only in appropriate patterns. To see something as having one belief requires seeing it as having many related beliefs (Davidson 1984, chap. 14,

200; Davidson 2001, chap. 7, 97–102; Stich 1983, 53–60). We would not attribute to someone the belief that a gun was in the desk drawer except insofar as we see her as believing that guns are artifacts, fire bullets, and have barrels, that desks are solid, that drawers open, have space in them, and so on. These general beliefs are conditions on possessing the concepts that are involved in the particular belief in question and express basic relations that hold between those concepts and other concepts and conditions relevant to their application. One must also typically have many beliefs about particulars to think a gun is in the desk drawer, which are supported by the general beliefs required to have the concepts, for example, that the desk is not alive, that it takes up space, that it is larger than the drawer, and so on.

Action, which is the expression of agency, is seen in the light of both belief and desire. Agents do things. We wave to friends, we write letters, we prove theorems, build houses, cross the street. For present purposes, we can remain neutral on what actions are. Candidates are bodily movements, construed broadly to include certain mental events (Davidson 1980, chap. 1), and causal processes (Dretske 1988). Actions are the products of intentions, which are formed in the light of our beliefs and desires. Typically the intention is formed on the basis of a desire for an end and a belief about how to achieve it. This shows the action in a favorable light, as done for a reason. Action explanations are often telescoped. We cite only the end or a connected means-end belief. We say, "He stepped on the brakes to stop the car," or "He thought she'd be impressed by flowers." The sense that an explanation has been given, however, depends on inferring that he thought stepping on the brakes would stop the car, or that he wanted to impress her. Davidson has argued influentially that every action is rationalizable by (made reasonable in the light of) a belief-desire pair that reflects means-end reasoning (Davidson 1980, chap. 1). This has been disputed on the grounds that some actions are done for their own sake, so that no belief that it conduces to a further end is required to provide its reason (Locke 1974, Mele 1988, Mele 2003b). In any case, each action can be represented as the correlate of the conclusion of a bit of practical reasoning about what it is best to do, and how to do it or what constitutes doing it. To see something as an agent then minimally requires seeing it as exhibiting coherent patterns of belief and a certain kind of reasonableness in acting.

While these observations show that some minimal level of coherence in thought, desire, and action is necessary for something to be an agent, they don't by themselves guarantee that an agent cannot have many inconsistent beliefs, reason mostly ineptly, or act mostly on reasons that are not best all things considered. Further support for the view that agents' attitudes as a whole must be seen as appearing in a largely rational pattern lies rather in reflection more generally on the conditions under which we are willing to treat something as an agent. We should expect some unclarity about where to leave off calling something

an agent. Like most natural language terms, "agent" is semantically vague. Our practice does not determine a precise cutoff point along dimensions of variation relevant to its application. However, the degree to which a system can be seen as rational is clearly a relevant dimension of variation for the term "agent." The less coherence we find in the set of attitudes we are thinking of as potentially those of an agent, the less clear we are that what we are considering is a possible agent at all. Moreover, reflection on cases—for example, step-by-step increases in overall incoherence—shows that to see another as an agent at all requires finding a large degree of coherence in his outlook, interests, intentions, and behavior. This coherence is expressed in seeing how the attitudes attributed both make for a reasonable picture of things, from the agent's point of view, and make sense of her behavior as an expression of agency. We do not treat anything as an agent unless we find a large degree of reason in what it thinks and does. This can escape our notice because we tend to focus on follies that we can identify only in the light of the vast background of reasonable action and belief that makes sense of people having the attitudes we take them to have. Irrationality, seen in this light, is a "perturbation of reason," not its absence (Davidson 2001, chap. 7, 99). We identify an irrational action or attitude as one that is a departure from an otherwise largely rational pattern, a pattern that makes sense of the attribution of attitudes that depart from it.

While this is the majority view, it has been challenged, nominally, at least, on both empirical and conceptual grounds. Investigation shows that these challenges, in addition to being subject to internal criticisms, involve a mischaracterization of the thesis.

Thus some psychologists have argued that experimental results show that human beings are not in fact rational animals (Johnson-Laird and Wason 1977, Nisbett and Ross 1980, Tversky and Kahneman 1983). "Pace Aristotle," one author says, "it can be argued that irrational behavior is the norm not the exception" (Sutherland 1994, vii). Here is an example. Experiments show that most college students do poorly on some versions of the selection task. Consider four cards, with "E," "C," "5," and "4" on their faces. Each card has a letter on one side and a number on the other. The task is to turn over the minimum number of cards to check whether, if a vowel is on one side, an odd number is on the other. Many subjects turn over the "E" card and the "5" card, but not the "4" card. The conditional is falsified if a card with a vowel on one side has an even number on the other, so the "E" and "4" cards should be turned over. These and other experiments have been alleged to show that *most* people reason *irrationally* on many *simple* reasoning tasks.

However, it is clear that these observations about mistakes on reasoning tasks (as understood by the experimenter) do not undermine the view that it is constitutive of the propositional attitudes that they occur in largely rational patterns. Identifying what subjects believe and want to do in the experimental situations

itself requires seeing their attitudes appearing in characteristically rational patterns. Their failures to reason as well as they could are themselves identifiable only because we already see them as largely reasonable creatures, responding largely reasonably to the tasks set before them. At most, then, these experiments could show there is some standard of rationality most people fail to meet. But, as Ernest Sosa (1999) has noted, to infer from this that most people are irrational or not rational is like inferring from the fact that human eyesight falls short of some extrahuman standard that most of us can't see well or can't see at all.

Apart from this, the interpretations of these sorts of experiments have been challenged on the grounds that they often involve overly simple assumptions about how the experimental subjects understand their task, and fail to distinguish between what subjects can do (competence) and what they do in the circumstances (performance). For example, we often use sentences to convey more than is conveyed by their literal meaning. We understand a card shark who says, "If it has this mark on the back, it's an ace," to be implying also that *all* aces have this mark on the back. Otherwise the remark would not be to the point. This would explain why in the selection task subjects turn over the "5" card. Inattention and nonsalience can help explain why many don't turn over the "4" card. Conditionals are typically used in *modus ponens* reasoning rather than *modus tollens* reasoning. Inattentive subjects then are apt to check inference potential first. However, they understand the mistake when it is explained to them. Thus, one must also distinguish between what one is capable of on reasoning tasks in principle and one's actual performance, which may be affected by a wide range of conditions. (See Cohen 1981 and Davidson 1980, chap. 14, 270–73, for further discussion.)

In addition to psychologists' efforts to show we are not in fact as rational as we might suppose, some philosophers have argued that there are only "minimal" conceptual constraints on the irrationality of agents. In particular, under this heading, it has been argued that there is no set of inferences or inference forms that all thinking beings must endorse, no " 'fixed bridgehead' of true and rational beliefs" (Hollis 1982, 73). The view that there is a fixed bridgehead of true and rational beliefs has been attacked in particular by Stephen Stich (1990, chap. 2), who follows Cherniak 1986 in rejecting it on the grounds that we can imagine a people whose feasibility ordering for inferences is inverted with respect to ours: inferences we find easy, they find hard, and vice versa.

Again, however, prima facie this is not a challenge to the thesis that propositional attitudes appear only in largely rational patterns. Nothing in the thought experiment suggests that the imagined people have massively inconsistent or incoherent beliefs or suffer from significant breakdowns of practical reasoning. In addition, the thought experiment that drives the argument is of doubtful coherence. The main difficulty is that the hypothetical others must possess the same concepts we do, for the complex inferences they are to find easy and we hard are

ones couched in terms of our concepts. But we do not attribute concepts to people who are not able to recognize the simplest of inferences they conceptually underwrite involving them. So anyone who possesses those concepts must be able to recognize the validity of the same simple inferences we do (Biro and Ludwig 1994).

3. Language and Rationality

Given that agency and thought are necessary for language, if rationality is a condition on agency and thought, it is a condition on possessing a language. This section discusses how rationality is related to interpreting others as speakers. The next section takes up the question of whether language is necessary for thought.

Rationality, Interpretation, and the Principle of Charity

Speakers are agents. Hence, to possess a language is to be at least as rational as agents must be in general. A constraint then on interpreting another as a speaker, on assigning meanings to his sentences, and contents to his attitudes, is that the pattern of assignments makes him largely rational. Moreover, the attitudes attributed must make sense of him as an agent capable of performing speech acts and communicating with others on a potentially limitless range of topics. Because seeing another as rational is a matter of finding his attitudes appropriately related by content and mode, this imposes a *holistic constraint* on interpretation, in the following sense. In finding patterns in another's behavior appropriate for interpretation as linguistic and other intentional behavior, one must be sensitive to the full pattern of assignments of meanings to sentences and attitudes in judging the appropriateness of the interpretation of any given bit of behavior.

Radical interpretation is interpretation of another without the usual contingent assumptions of commonality of language, culture, or psychology (Davidson 1984, chap. 9). Reflection on radical interpretation is a tool in investigating the most general and abstract requirements on having a language. When we strip away all the common aids to interpretation and consider how we could come to interpret another simply on the basis of whatever we can know a priori about speakers, and on behavioral evidence open to public observation, we uncover what patterns in behavior we must perforce think are there if it is to be interpretable as linguistic behavior. This helps to show what content our linguistic and allied concepts

have—those of meaning, truth, and the propositional attitudes—by showing how they are related to independent evidence for their application.

The requirement that we see others as largely rational in order to interpret them has most famously been discussed as a constraint on radical interpretation under the heading of "The Principle of Charity." We will concentrate here on Davidson's version of the principle. The principle has two distinguishable aspects, which are motivated differently. In his later work, Davidson has distinguished the two strands as the Principle of Coherence and the Principle of Correspondence (Davidson 1985a, 92; Davidson 2001, chap. 14, 211).

The Principle of Coherence is concerned specifically with the a priori requirements on seeing something as an agent that can perform speech acts. It enjoins one to find another to be largely epistemically and practically rational, on pain simply of not being able to see the other as an agent at all. The grounding for this part of the Principle of Charity is a priori reflection on the nature of agency and the propositional attitudes of the sort sketched briefly in the previous section of this chapter. Davidson does not attempt to spell out in general the requirements on seeing others as largely rational, but argues that in practice it is a matter of seeing another as rational and reasonable by one's own lights, adjusting for differences in position and interests. This is consonant with seeing the principle as appealing to a priori constraints. For to say something is rational by our lights is just to say that we see it as required of a rational agent given our concept, that is, *the* concept, of rationality. Thinking about rationality from the point of view of interpreting another (or seeing another as an agent), in the light of the other's behavior, sheds some additional light on the concepts of agency and of the propositional attitudes, for it helps to highlight what gives them their practical content, namely, the way in which they enable us to make systematic sense of the movements of an object, by seeing it as goal directed in the light of largely true beliefs about the environment. To the extent to which we *cannot* make good sense of a system in these terms, we should not see it as being an agent with propositional attitudes: the supposition that it has propositional attitudes in this case is *idle*, and disconnected from the role the concepts of the attitudes play in our making sense of things. (See Davidson's work in the philosophy of action, in particular the essays collected in Davidson 1980, esp. chaps. 1–3, 5, 12–14; also Davidson 1982 and Davidson 2001, chap. 7.)

The Principle of Correspondence has received three different formulations. To explain these and the motivation for the principle, it will be necessary to say more about Davidson's account of radical interpretation, for the Principle of Correspondence is introduced to solve a problem that faces the radical interpreter when his task is cast in a certain way. In particular, Davidson gives a central place in the interpreter's procedure to confirming for a speaker's language a formal recursive truth theory similar to the sort characterized by Tarski (1932) 1983. The

truth theory is to serve as the vehicle for a compositional meaning theory for the language (Ludwig 2002). The theory has as theorems sentences of the form (T) (or notational variants)

(*T*) *s* is true-in-*L* iff *p*

where "*s*" is replaced by a description of a sentence of the speaker's language in terms of its primitive significant parts, and "*p*" by a sentence in the interpreter's language. The radical interpreter aims to construct and confirm a truth theory that issues in theorems of this form by observing the speaker in his environment. Davidson has suggested that a truth theory confirmed from this standpoint can be used to interpret the speaker's utterances, that is, that we can use the sentence "*p*" used in the interpreter's language to give truth conditions for the speaker's sentence *s* in (*T*) to interpret *s*. This is in effect to hold that if a truth theory is confirmed from the radical interpreter's standpoint, "is true-in-*L* iff" in (*T*) can be replaced with "means that." (For natural languages, which contain context sensitive expressions, such as demonstratives and tensed verbs, the truth and meaning predicates must be relativized to features of context relevant to their interpretation.)

The radical interpreter must determine the appropriate (*T*)-sentences for the speaker's language. Davidson supposes the radical interpreter can, from behavioral evidence, figure out what sentences of his language a speaker thinks are true (holds true). The radical interpreter can then identify the conditions in the environment in which the speaker holds true a sentence. For some sentences, whether a speaker holds them true or not will not be sensitive to what goes on in his environment (e.g., "$2+2=4$"). In other cases, there will be systematic correlations between the speaker's holding true a sentence and what's happening in his environment (as in the case of "That's a rabbit"). These latter are the key to what the speaker means by his words and what he thinks.

A speaker's holding true a sentence is a "vector of two forces": what he believes and what he means by the sentence (Davidson 1984, chap. 10, 196). Specifically, Davidson assumes (idealizing somewhat) that if a speaker believes that *p*, and believes that a sentence *s* of his expresses that *p* (at the time), then the speaker infers that *s* is true. For the speaker is committed to its being true that *p* and is presumed to know that if it is true that *p*, and *s* expresses that *p*, then *s* is true. Accepting this, if we knew which sentences a speaker held true and what they meant, we could infer what he believed. Likewise, if we knew what the beliefs were that formed the basis for his holding true various sentences, we could infer what the sentences meant. We start out, however, knowing neither what he believes nor what his sentences mean. The role of the Principle of Correspondence is to show how to break into this circle.

The first of the three formulations the Principle of Correspondence has received is that other speakers are largely in *agreement* with one, explicable error and ignorance aside. The second is that when a speaker holds true a sentence, by and large the sentence is true. The third is that a speaker's *beliefs*, particularly those that are responses to his environment, are largely true. Each has some textual support. However, the third is the correct formulation, for that is the only one that points to a way of fixing either what someone believes or the meanings of his sentences. For if we can assume that a speaker's beliefs about his environment are mostly true, then we know that when his beliefs are correlated with things in the environment, his beliefs are likely to be about things they are correlated with. If we further assume that there are enough constraints overall on interpretation to narrow down to a single salient condition, for each belief, what it can be reasonably correlated with for interpretation, we can identify what he believes on the basis of the correlated conditions that prompt his beliefs, for that is what they are about. If the third interpretation of the principle is correct, the correctness of the first two follows, but not vice versa.

The Principle of Correspondence is justified as a necessary condition on being able to construct a justified interpretation of another speaker on the basis of (i) a priori constraints imposed by our conceiving of him as an agent possessing a language, and (ii) information about his interaction with his environment. It plays a role similar to the Principle of the Uniformity of Nature, which holds that nature evolves in accordance with general laws, in Hume's account of what is necessary to justify our inductive inferences. We infer from past regularities to future regularities. For this to be reasonable, we must assume minimally that nature evolves according to general laws, which are reflected in past observed regularities, which can then be projected into the future.

Suppose that we knew that we could be justified in believing things about the future on the basis of the past. Then we would be able to infer that it is reasonable to accept the Principle of the Uniformity of Nature, since if it were not, we could not be justified in our inductive inferences. The Principle of Correspondence is justified in a manner similar to this, by appeal to the assumption that we can come to a justified interpretation of another speaker on the basis of the evidence and a priori constraints available in radical interpretation, and the claim that if we were not justified in believing the Principle of Correspondence, we could not come to a justified interpretation of another speaker on the basis of the available evidence, given the constraints.

The argument to justify the Principle of the Uniformity of Nature by appeal to our being justified in making inductive inferences is question begging. For we cannot know our inductive practices are reliable a priori, and any a posteriori justification of inductive practices must rely on the Principle of the Uniformity of Nature. Hume argued that the Principle of the Uniformity of Nature, since it

is itself a generalization that covers the past and future, and is not knowable a priori, could be justified only by appeal to induction, thus leaving both the principle and our inductive practices unjustified.

For the Principle of Correspondence to fare better than the Principle of the Uniformity of Nature, we must be able to know a priori that we can arrive at correct interpretations of any other speaker on the basis of radical interpretation. If we could know it only a posteriori, and the assumption of the Principle of Correspondence were necessary in any inference to what a speaker means by his words and what he believes, we could not support the Principle of Correspondence in this way without presupposing it.

Davidson adopts the strategy of justifying the Principle of Correspondence by appealing to an a priori justification of the possibility of correctly interpreting any other speaker from the standpoint of the radical interpreter. The argument for the possibility of correctly interpreting any other speaker from the standpoint of the radical interpreter rests on the observation that language is by its nature a medium for communication, so that interpretation must be something that can be accomplished on the basis of evidence that is available interpersonally: "That meanings are decipherable is not a matter of luck; public availability is a constitutive aspect of language" (Davidson 1990, 314). The central idea here is that it is of the nature of a language that it is a device that enables its speakers to communicate with others, and that this can be done only on the basis of interpersonal evidence for what others mean.

Yet granting this does not seem to be enough to yield the result that Davidson needs. It would seem enough to satisfy the requirement that language be a medium for communication that it guarantee that *if* we had mostly true beliefs *and* were confronted with someone who speaks a language with largely overlapping expressive powers, *and* who believed similar things about the environment, we would be able to interpret him. Justifying the Principle of Correspondence requires the stronger claim that to have a language is to be interpretable correctly on the basis of public evidence in any circumstances by any possible speaker (see Lepore and Ludwig 2004, Ludwig 1992, and Ludwig 1999 for further discussion).

Even if we despair of justifying the Principle of Correspondence, we need not despair of being able to interpret others correctly. For we need not suppose that our epistemic resources are restricted to those available to the radical interpreter. We can appeal to knowledge of features of our own psychological type and to the fact that in practice others whom we want to interpret are conspecifics embodied in the same way we are and in similar environments, to infer with some plausibility the sorts of things they are apt to be thinking, in order to constrain our interpretations. This would not, however, yield any argument to the conclusion that the empirical beliefs of linguistic beings were by their nature largely correct.

What is the relation between the two parts of the Principle of Charity? Traditionally, being rational has not been seen as requiring largely true empirical

beliefs. Thus, it has been supposed that one could be rational but mostly mistaken in one's empirical beliefs, someone, for example, who reasons perfectly, but who is systematically deceived about his environment by Descartes's Evil Demon, and, so, through no fault of his own. Rationality on this view is a precondition on having largely true empirical beliefs but does not require it. Rationality would require a large number of true general beliefs as a condition on possessing the concepts involved in any of an agent's beliefs, but these would not be empirical. For example, to possess the concept of red, one would have to believe, indeed, to know, that red is a color, that red is a feature of the surface of an object, that no surface, viewed from one position, can be two different colors at the same time, that surfaces are extended, that extended objects occupy space, and so on. These propositions are not empirical propositions; they are necessary, and knowable a priori. If Davidson is right, one cannot be rational and a speaker without having mostly true empirical beliefs as well. To be a rational speaker, then, would be to be largely right about the world. Yet, while being largely right in one's empirical beliefs, if Davidson were right, would be necessary for being a rational speaker, it would not thereby be an aspect of being rational.

I have characterized the Principle of Coherence as being grounded in a priori reflection on the nature of agents. However, thinking about its role specifically in the light of interpretation of other speakers is useful in seeing why this is so. For if we think about our application of the concepts of the attitudes to other agents, it is clear that their utility lies in their enabling us to discern a pattern in the behavior of others that is usefully projectible. As Daniel Dennett (1987) puts it, adopting "the intentional stance" enables us to explain and predict the behavior of complicated systems in a way that would otherwise be practically impossible. Thus, contrast trying to predict where someone's body will be tomorrow at noon on the basis of its physical constitution and the laws of nature, on the one hand, with trying to predict where it will be on the basis of his intending to keep a lunch appointment with you tomorrow at the faculty club, on the other. It is the patterns among the attitudes and their relations to behavior imposed by the requirement that agents be largely rational that make for the practical utility of descriptions of objects in terms of propositional attitudes. That is, if a system has successfully been understood as an agent (and, hence, as rational) in the past, so that we have succeeded in interpreting its behavior as the product of rational agency, then, on the assumption that it continues to be an agent, we can usefully (if roughly) predict its future behavior from its present attitudes, any new attitudes it acquires, and what we suppose its reasoning powers to be. (Carl Hempel argued that it is the empirical assumption of rationality that enables us to predict agents' future behavior [Fetzer and Hempel 2001]; if being largely rational is essential to agency, the relevant empirical assumption is that the system is an agent. Of course, assumptions about the *degree* of an agent's rationality can still play a role in empirical explanation. In this regard, see also Davidson 1980, chap. 14.)

Language as Necessary for Rationality but Not for Thought

The discussion has assumed that rationality is constitutive of thought, and, hence, that it is constitutive of language. But, as noted in the introduction, it has been argued that while language is necessary for rationality, rationality is not necessary for thought. On the face of it, both views cannot be correct. However, if we consider what has been said in favor of the view that language is required for rationality, though rationality is not required for thought, we will see that proponents of the view have in mind stronger requirements on rationality than those we have imposed above. Thus, these two views, though prima facie in conflict, are not actually so.

The view that language is necessary for rationality, but that rationality is not necessary for thought, has an ancient pedigree. Aristotle, and the Stoics, following him, held that while nonlinguistic animals had mental capacities, they did not have the capacity for reason, or, consequently, for belief (Sorabji 1996). This view was grounded on two claims. The first is that belief is assent or conviction as a result of persuasion (by others or oneself). The second is that reason is absent where there is no capacity for persuasion. It can be seen, however, that this is not so much a denial of anything we have said above as a use of a stronger conception of "belief" (*doxa*) or "reason" (*logos*) than we have been considering. For this is to think of *doxa* as what we would call belief arrived at on the basis of reflective reasoning, and *logos* as the capacity for reflective reasoning.

This brings out an important point about how "rationality" is now predominantly understood: not as a matter in the first instance of the capacity to engage in reflective reasoning but rather as a matter of thought and action being seen as reasonable in the light of evidence, belief, and desire, whether the agent reflects on these or not. Being rational, in the sense characterized in the previous section, and having the power of rational *reflection* are not the same.

More recently, Jonathan Bennett (1989, 93) has argued that "possession of language is necessary for rationality" (though not sufficient), on the grounds that

(1) "Rationality requires the ability to manifest in behavior judgments about what is particular and past and what is general, that is, to manifest behavior that is appropriate or inappropriate to that which is not both particular and present";

(2) "Only linguistic behavior can be appropriate or inappropriate to that which is not both particular and present" (87).

Bennett explicitly denies, however, that "rationality" as it appears in this claim and argument "has anything to do with 'rationality' in any contemporary sense of that term" (viii), and is rather intended "to stand for whatever human possession it is that creates a mentalistic difference of kind between us and other terrestrial animals" (vii, 4–5), that is, with giving a sense to "rational" that would

truly make us the only rational animals. It is not surprising, given this stipulation, that it should turn out that only linguistic beings are rational. This thesis, then, like Aristotle's, and others in a similar vein, hinges on a conception of rationality that is not the one at issue above in the claim that rationality is necessary for any thought.

4. Language and Thought

The final question we take up is the relation between language and thought. If thought is prior to language in the sense that there can be thinking beings who are not linguistic beings, then the rationality of speakers is inherited (in part) from the rationality required of agents. If language is required for thought, on the other hand, then there is a more intimate connection between having the power of speech and being rational. For then the power of rational thought and the power of speech would be interlocking capacities, each required for the other.

The majority view, unsurprisingly, is that thought does not require language. It is easy to see why this is so. First, it is natural and effective to adopt the intentional stance toward many animals. We explain the behavior of both do-mestic and feral animals by attributing to them beliefs about the world around them and desires similar to ours in basic respects. As Hume says, "When . . . we see other creatures, in millions of instances, perform like actions [to ours], and direct them to like ends, all our principles of reason and probability carry us with an invincible force to believe the existence of like causes" (Hume [1739] 1978, 176). Second, we suppose ourselves to be continuous with the rest of the natural world, and it can seem incredible that human beings should represent the only evolved animals capable of even rudimentary thought. This seems especially incredible in the light of the natural view that language is a means of expressing thoughts that we antecedently possess. This idea is clearly expressed in book 3 of Locke's *Essay Concerning Human Understanding* ([1690b] 1987), which has had a profound effect on the development of thinking about the mind, even though it has long been recognized that Locke's theory is inadequate in its details. Given the ease with which we can see the nonlinguistic behavior of other animals as expressive of the same sorts of things that similar behavior in us expresses, the view that language merely adds the capacity to express the same basic types of psychological state can seem compelling.

There have been dissenters, however, from ancient times to the present, in different forms. Aristotle, as we've seen, held that animals were incapable of belief without language (though he seems to have had in mind more by *doxa* than we

would require of belief). Plato held this too, though by virtue of holding implausibly that thought is silent speech; since Plato held that animals have thought, he attributed to them the power of language as well. Descartes famously denied nonlinguistic animals had minds (Descartes 1984a, 140–41; Descartes 1984b, 302–3; Malcolm [1973] 1991). Kant thought only the fully rational could have beliefs, and only those who possessed language could be fully rational. In contemporary philosophy, the independence of thought from language has been most famously challenged by Donald Davidson (Davidson 1984, chap. 11; Davidson 2001, chap. 7).

Davidson gives two arguments for the claim that only speakers have thoughts. The first is the argument from holism (cf. Stich 1979). The argument focuses on the kind of evidence available for the attribution of attitudes in the absence of verbal behavior.

The Argument from Holism

- (A1) Beliefs and other attitudes are ascribable only in dense networks of attitudes.
- (A2) Attributing a dense network of attitudes to an agent requires for support a rich pattern of behavior that gives substance to the attributions.
- (A3) The pattern of behavior required cannot be sustained in the absence of verbal behavior.
- (A4) Therefore, only linguistic animals can have propositional attitudes.

Davidson recognizes that this argument is not conclusive. The main problem lies in sustaining premise (A3). The support comes from recognizing that there is something arbitrary about how we choose to describe the supposed attitudes of nonlinguistic animals. As Davidson says,

> If we really can intelligibly ascribe single beliefs to a dog, we must be able to imagine how we would decide whether the dog has many other beliefs of the kind necessary for making sense of the first. It seems to me that no matter where we start, we very soon come to beliefs such that we have no idea at all how to tell whether a dog has them, and yet such that, without them, our confident first attribution looks shaky. (Davidson 2001, chap. 7, 98)

> A very complex pattern of behavior must be observed to justify the attribution of a single thought. . . . I think there is such a pattern only if the agent has language. (100)

Thus, what is there to choose between saying that Rover thinks *a cat went up that tree* or saying that Rover thinks *a furred pointy-eared quadruped with whiskers*

scurried up the oak he's barking up? Is there any further behavior that Rover could display that would decide which of these was more appropriate? It seems clear we should not take the particular sentence chosen for the belief attribution too seriously. How then can we take seriously the idea that Rover has any beliefs at all, for if he does have beliefs, must they not have determinate contents?

It is undeniable, however, that there is a pattern of activity that is captured equally well for practical purposes by either of our two attributions, but not by attributing to Rover the belief that the cat has vanished or turned into an acorn: Rover barks up the tree, peers into the branches, runs around its base, and finally settles down to wait (as we put it). (And there is the tale of Chrysippus's hunting dog, who, faced with three roads, checks the scent down two, before setting off down the last without checking.)

In light of this, perhaps we can characterize Rover's attitudes with classes of sentences that would do equally well for the purposes of accounting for his non-verbal behavior. Each particular sentence in the class attributes too much conceptual sophistication to Rover. To characterize his belief as captured equally well by any of them is, however, to characterize exactly the determinate content of his belief and the degree of refinement of his concepts, though we do not have words to express this in a single sentence. We can press into service Davidson's own favored analogy for disarming the threat of the indeterminacy of interpretation (Davidson 1984, chap. 10, 154). Like the different scales we could use in measuring temperature, the different belief sentences we use to keep track of Rover's behavior (each of course requiring a supporting pattern of related ascriptions) all equally well capture the phenomenon we are interested in, which is different from saying that there isn't any (Jeffrey 1985).

Davidson's primary argument, though, rests on claims about what is required to have beliefs, namely, the concept of belief, and what is, in turn, required to have the concept of belief, namely, language. The basic argument is as follows (Davidson 1984, chap. 11, 170; Davidson 2001, chap. 7, 102).

The Argument from the Concept of Belief

(B1) One can have propositional attitudes (thoughts) only if one has beliefs.

(B2) One can have beliefs only if one has the concept of belief.

(B3) One can have the concept of belief only if one has a language.

(B4) Therefore, one can have propositional attitudes (thoughts) only if one has a language.

The crucial premises are (B2) and (B3). Davidson offers an argument for (B2) in "Rational Animals" (Davidson 2001, chap. 7, 104).

The Argument from Surprise

(C1) "One cannot have a general stock of beliefs of the sort necessary for having any beliefs at all without being subject to surprises that involve beliefs about the correctness of one's own beliefs. Surprise about some things is a necessary and sufficient condition of thought in general."

(C2) "Surprise requires the concept of belief."

(C3) Therefore, one has beliefs only if one has the concept of belief.

(C2) seems true. Surprise requires recognition that something one thought was so is not, and so requires that one have the concept of belief. (C1) appears more susceptible to challenge, however. Davidson appeals bluntly to intuition, but this seems unconvincing. Child psychologists have argued that children (after initial language acquisition) pass through a developmental stage in which they cannot recognize that they had false beliefs (Gopnik 1990; Gopnik and Astington 1988; Perner, Leekam, and Wimmer 1987). If so, then despite possessing a language, and so thought, they are not capable of being surprised. Whether this is right or not, it certainly seems possible. But if so, there can be no a priori requirement on thought or language that one be capable of being surprised. (C1) would furthermore lead to the surprising conclusion that there cannot be an omniscient being, since such a being is never surprised. However, we can secure (C3) by another route. For, granting that belief requires agency, having beliefs requires having beliefs about actions serving as means to ends (a belief that doing *this* is likely to bring about *that*), and so the concept of action. But the concept of an action involves in turn the concepts of belief, desire, and intention. Thus, one can have beliefs at all only if one has the concept of belief. Let us therefore grant (B2).

That leaves (B3). Davidson's main argument for (B3) can be reconstructed as follows (see Davidson 1984, chap. 11; Davidson 2001, chap. 3, chap. 7, 103–5, chap. 13, 202 for (D3), chap. 14):

The Argument from the Concept of Error

(D1) To have the concept of a belief, one must have the concept of error, or, what is the same thing, of objective truth, the contrast between how things are represented and how they are.

(D2) The claim that a creature possesses the concept of error, or objective truth, stands in need of grounding: this must take the form of giving an account of how there could be scope in the creature's experience for correct application of the concept.

(D3) We can understand how there could be scope for the application of
the concept of error in a creature's experience if, and only if, we con-
ceive of it as a creature that is, has been, or is potentially (fixing the
creature's capabilities) in communication with other creatures, and so
able to use the concept of error as a tool in interpretation to achieve
a better rational fit between a speaker's behavior and the beliefs and
meanings we attribute to him; that is, the concept of error would
have some work to do for interpreters of others' speech (there would
be scope for its application in their practice), but not otherwise.

(D4) Therefore, from (D2)–(D3), to have the concept of error or objective
truth one must be, or have been, or potentially be in communication
with others.

(D5) Therefore, from (D1) and (D4), to have the concept of belief, one
must be, have been in, or potentially be in communication with others.

Since communication with others requires one to have a language, (D5) suffices
to establish (B3). Davidson nowhere states (D2) explicitly, but it appears to be in
the background in the argument, which proceeds by asking after conditions that
make sense of there being scope for the application of error in a creature's be-
havior or experience. Davidson states (D3) in a stronger form than given here, as
requiring actual communication, but this seems too strong and is not required
for the conclusion.

(D2) may be challenged on the grounds that all that is required for having a
concept is that a creature has reason to think that there is scope for correct
application of it in his experience. This is a familiar difficulty in transcendental
arguments that rely on conditions for the possession of concepts to get to their
correct application. However, even granting (D2), (D3) seems questionable. For
while communication provides scope for the application of the concept of error,
it is not clear why many other activities do not provide just as much scope for
its application, such as, for example, correcting one's own past beliefs in the light
of new evidence, or explaining behavior of a nonlinguistic animal that otherwise
seems irrational in the light of a false belief, for example, explaining Rover's
barking up the wrong tree by attributing to him the false belief that that is where
the cat is.

Davidson's argument for the claim that language is essential to thought is
therefore inconclusive. Consequently, we cannot support premise (2) in the fol-
lowing argument for the conclusion that language is necessary for rationality by
way of Davidson's argument: (1) a being can think iff it is largely rational; (2) a
being can think iff it possesses the power of speech; therefore, (3) a being is largely
rational iff it possesses the power of speech. (See Heil 1992, chap. 5, and Lepore
and Ludwig 2004, chap. 22, for further discussion.)

5. SUMMARY

Rationality is essential for thought and agency. It is constitutive of the propositional attitudes that they are attitudes of an agent, and nothing is an agent except insofar as its behavior can be interpretable as expressing largely rational behavior and thought. Since agency is a condition on language, rationality is essential for language as well. To say that thought requires rationality is not to say thinkers, and speakers, are perfectly rational, incapable of mistakes, or follies, neuroses, psychoses, and so on. Rather, what mistakes agents make, what irrational behaviors they engage in, are identifiable only against a background of largely rational thought and behavior, which makes sense of the attitudes involved in the irrational behavior or reasoning. Language may be argued to be essential for rationality only by employing a special standard for rationality, requiring advanced reasoning abilities, or by arguing that one can have propositional attitudes only if one is a speaker. Davidson's controversial arguments for the latter position are not conclusive. The ease with which we can explain animal behavior by attributions of propositional attitudes on the basis of the same sort of nonverbal behavior we often use with human beings suggests that thought, agency, and, hence, rationality, are not confined to linguistic beings. Language is not a precondition of rationality, but rather amplifies our powers of reasoning.