Thought and language

Michael Pendlebury

Department of Philosophy
University of the Witwatersrand, Johannesburg
Private Bag 3
Wits, Johannesburg
2050 South Africa
E-mail: <103mjp@muse.wits.ac.za>

Abstract
This article defends the view that nonlinguistic animals could be capable of thought (in the sense in which the mere possession of beliefs and desires is sufficient for thought). It is easy to identify flaws in Davidson’s arguments for the thesis that thought depends upon language if one is open to the idea that some nonlinguistic animals have beliefs. It is, however, necessary to do more than this if one wishes to engage with the deeper challenge underlying Davidson’s reasoning, viz., that of providing a principled account of what it takes for a representer to qualify as a thinker. Heil attempts to construct a Davidsonian account on the basis of the hypothesis that the semantic opacity essential to thought is rooted in second-order representation (which Davidson ties to language), but it can be shown that second-order representation is neither necessary nor sufficient for opacity. A reasonable non-Davidsonian account of thought in terms of which sufficiently sophisticated nonlinguistic animals qualify as thinkers is, however, possible.

1. Introduction

We may reasonably take it for granted that language depends upon thought inasmuch as a being that does not think could not be a competent user of a language. I will be concerned with the more contentious question of whether thought also depends upon language – in the sense in which this would rule out the possibility of a thinking being that is not linguistically competent. This dependence thesis is not the strongest imaginable, for (providing the thinker is linguistically competent) it does not exclude the possibility of particular thoughts that are not linguistic or could not be articulated linguistically. The main purpose of this article is to present the case for the view that nonlinguistic animals could be capable of thought in a way that does reasonable justice to the arguments and concerns of those who oppose this position. But this will not be possible without a review of some well-trodden territory.

I shall use the term ‘language’ to cover public natural languages, like English, Tswana or Gujarati, as well as any actual or possible communication systems that are to some degree like such languages with respect to the complexity of their syntactic and semantic structures, and the richness and adaptability of their expressive powers.

1 For arguments in support of the claim that some nonlinguistic animals have thoughts that are more direct and less accommodating to the opposition, see, e.g., Wilkes 1997 and Weiskrantz 1997.
Although this is not very precise, it is easy enough to apply. For example: the English of normal 4-year-olds counts as a language in these terms; the natural signalling systems of honey bees and vervet monkeys, and the artificial symbol systems taught to some captive chimpanzees, do not qualify as languages; and the time at which the proto-language of the developing child becomes language proper is left indeterminate. Greater precision is not required for anything that follows.

To the best of our knowledge human beings are the only animals that possess language as I have characterised it. Thus those who hold that thought depends upon language are implicitly denying that nonhuman animals can think. This view has been endorsed by various philosophers since ancient times (see, e.g., Aristotle, De Anima, Book III), and it is easy to accept if, along with Dennett (e.g., 1996, p.130), we restrict the term ‘thinking’ to manifestations of what Kant called the faculty of understanding – in other words, to mental states, events and processes like those involved in the kind of far-reaching, self-conscious reasoning and deliberation that is characteristic of human beings. If thinking is thus restricted to reflective reasoning and judgment, then it is plausible not only that non-human animals do not think, but also that (as a matter of empirical fact) human beings cannot realise their capacity for thought in the absence of language.

However, most of us who hold that some nonlinguistic animals, including higher mammals, can think, do not mean to ascribe full-blown reasoning and judgment to them. A broader understanding of thought is, therefore, required if we are to join the debate. Fortunately, we are able to accept the proposal by Donald Davidson, the outstanding champion of the view that thought depends upon language, that in order for a being to qualify as having thoughts it is sufficient for it to have beliefs (see Davidson 1975/1984, pp.156-7 and 1982/1985, pp.473-4).

It is, however, unfortunate that Davidson adopts the widespread practice in analytical philosophy of characterising beliefs, desires and the like as propositional attitudes. The problem with this is that there is a strong tendency for philosophers to view propositions as more-or-less linguistic even when this is not intended, and the label ‘propositional attitude’ is therefore apt to suggest that there is something linguistic about beliefs. I suspect that Davidson is taken in by this suggestion, and his talk about the propositional contents of beliefs is never quite open to the idea that those contents could be thoroughly nonlinguistic.

This concern is exacerbated by Davidson’s characterisation of beliefs and desires as reasons, and of explanations of behaviour in terms of beliefs and desires as rationalis-

---

2 For a little more on the nature of such thinking see the discussion of the difference between judgment and belief in Pendlebury 1999, pp.361-3.

3 Merely possession of the human ‘faculty of language’ (see, e.g., Chomsky 2000, especially pp.168-72) may suffice for some distinctively human cognitive powers, but full-blown reasoning and judgment no doubt require the realisation of the capacities involved in that faculty through the mastery of a public natural language.


5 I myself have argued that the contents of perceptual experiences are propositional, meaning only that they represent whole states of affairs, and that they are therefore subject to truth and falsity (Pendlebury 1990), but I find that readers often suppose that I must have meant to suggest that those contents are somehow linguistic – even though I took pains to make it clear that this was not my view.
ations (see, e.g., Davidson 1963/1980, pp.3-4, and 1975/1984, pp.159-60); by his proposal that having some propositional attitudes is all that it takes for a being to qualify as a rational animal (see Davidson 1982/1985, p.473); and by his view that animals like dogs do not have beliefs because “They do not form judgments” (Davidson 1999, p.8 – emphasis added). All this begs the question against the possibility of beings that have thoughts without being capable of reasoning, deliberation and judgment. Thus, while we may agree with Davidson that belief is sufficient for thought, we should be on our guard about his views on what belief involves and what follows in its wake.

2. Davidson’s Case Against Thought Without Language

Davidson allows that we are pragmatically justified in attributing beliefs and desires to nonlinguistic animals on the ground that this is our best way of explaining and predicting their behaviour, but he nonetheless insists that such animals do not really have beliefs and desires. For when we attribute such states to them ‘we are applying a pattern of explanation that is far stronger than the observed behaviour requires, and to which the observed behaviour is not subtle enough to give point.’ (Davidson 1982/1985, p.478; see also Brandom 1994, p.155.) In Dennett’s terminology, Davidson’s position is that we have no option but to adopt ‘the intentional stance’ (see, e.g., Dennett 1996, pp.27-40) towards some nonlinguistic animals even though they do not literally have the intentional states, like belief and desire, that we ascribe to them.

Dennett, in contrast, takes the line that a being’s having intentional states is just a matter of its being useful to adopt the intentional stance towards it (Dennett 1987, Chapter 2 and 1996, p.34). Thus, although he holds that only linguistic animals are capable of thinking, ‘Dennett ... insists that ascriptions of beliefs and desires to single-celled organisms, plants and artifacts are no more metaphorical than is the ascription of beliefs and desires to our fellow human beings’ (Heil 1998, p.155). Although I will not discuss Dennett’s position in its own right, my reasoning in this article will cast considerable doubt upon it.

Davidson’s case for the view that a being without language could not actually have beliefs depends on association and suggestion as well as explicit argument, and it also involves appeals to aspects of his philosophy on which the issue should not have to turn, for example, his paratactic account of attributions of propositional attitudes and his theory of radical interpretation. Thus any presentation of Davidson’s case is likely to leave out something of significance. This applies especially to the following brief sketch, which seeks to uncouple the main lines of Davidson’s arguments from his other philosophical commitments.

I see Davidson’s case as a campaign on two fronts. On the first front he tries to establish that for various connected reasons we could never be justified in attributing a

6 The word ‘intentional’ and its cognates are sometimes used very broadly, to cover all phenomena exhibiting some form of ‘directedness’, however weak (including the heliotropism of sunflowers and the so-called representational states of ordinary thermostats), and sometimes more narrowly, to cover only the most sophisticated forms of directedness, most notably those displayed by language and mental states and events such as beliefs, desires and occurrent thoughts. In this article I adopt the more restricted use.

7 On the basis of Dennett 1991a, alternative readings of Dennett are, however, possible. I owe this point to David Spurrett.

8 More systematic and expansive presentations of Davidson’s case appear in, e.g., Chapter 6 of Heil 1992 and Glock 2000. I consider Heil’s most important contributions to the debate in sections 4 and 5 below.
genuine belief to a nonlinguistic being. As I am inclined to cut them up, the most important of these reasons are as follows.

First, we cannot assign beliefs to an animal independently of its desires, and whenever we can assign ‘one constellation of beliefs and desires’ to the animal on the ground that they would explain its behaviour, ‘it is always possible to find a quite different constellation that will do as well’ (Davidson 1975/1984, p.160). Thus the attribution of beliefs and desires on the basis of nonlinguistic behaviour alone leaves them completely indeterminate. This indeterminacy is avoidable only if linguistic behaviour is taken into account and the attribution of beliefs and desires goes ‘hand in hand with the interpretation of speech’ (Davidson 1975/1984, p.163).

Second, any single belief is always part of an extensive belief system, and its content and identity depend on its place within that system. As Davidson puts it, ‘In order to believe that the cat went up the oak tree I must have many true beliefs about cats and trees, this cat and this tree, the place, appearance and habits of cats and trees, and so on...’ (Davidson 1982/1985, p.475). Thus one cannot be fully justified in attributing a given belief to a being unless one is also entitled to attribute numerous other related beliefs to it. But in the case of a nonlinguistic animal like a dog, ‘It seems that no matter where we start, we very soon come to beliefs such that we have no idea at all whether a dog has them, and yet such that, without them, our confident first attribution looks shaky’ (Davidson 1982/1985, p.475).

Third, beliefs have fine-grained contents which makes them subject to semantic opacity, so that sentences used to report their contents are liable to substitution failure. Suppose that

1. Pete believes that the first Postmaster General of the United States was a statesman

is true. Even though the terms ‘the first Postmaster General of the United States’ and ‘the inventor of bifocals’ denote the same man, namely, Benjamin Franklin, the sentence which results when we substitute the latter for the former in (1), namely,

2. Pete believes that the inventor of bifocals was a statesman

could easily be false.

If nonlinguistic animals really do have beliefs, then (given a few bells and whistles) sentences reporting those beliefs should also be liable to substitution failure. But it is not clear that this is the case:

9 A common variation of this argument claims that all beliefs require concepts, which are possible only within a system of general beliefs that cannot justifiably be attributed to nonlinguistic animals (see, e.g., Glock 2000, pp.42-4). In the body of this article I avoid the issue of whether nonlinguistic animals have concepts because it introduces unnecessary terminological complications. I am inclined to adopt the Kantian practice of restricting the term ‘concepts’ to possible constituents of judgments, but more latitudinarian uses of the term are common (see, e.g., Allen 1999 and Glock 2000, pp.42-9) and legitimate. Nonlinguistic animals probably do not have concepts in my demanding sense, but it is not obvious that concepts thus understood are required for beliefs.

10 I depart from standard practice insofar as I apply the phrase ‘semantically opaque’ directly to beliefs and other representations rather than to sentences used to report their contents. This is because our interest is in a possible property of representations in their own right, and because I will claim in section 5 that substitution failure in sentences reporting their contents is not always an adequate test of the presence of that property.
[H]ow about the dog’s supposed belief that the cat went up that oak tree? That oak tree, as it happens, is the oldest tree in sight. Does the dog think that the cat went up the oldest tree in sight? Or that the cat went up the same tree it went up the last time the dog chased it? It is hard to make sense of the questions. But then it does not seem possible to distinguish between quite different things the dog might be said to believe (Davidson 1982/1985, p.474).

If so, then it seems that we cannot be justified in attributing any particular belief to the dog.

These three first-front arguments are linked by what Davidson describes as ‘the holism of the mental realm’ (Davidson 1970/1980, p.217), in terms of which the identity and behavioural effects of particular intentional mental states of a subject – including its beliefs and desires – depend upon the rest of its intentional system.

On his second front Davidson offers a single two-stage argument for the conclusion that only linguistic beings can have beliefs.

First, any believer must have the concept of belief, because ‘Someone couldn’t have a belief unless he understands the possibility of being mistaken’ (Davidson 1975/1984, p.170), and because believers must be capable of surprise when they are mistaken (Davidson 1982/1985, p.479). Both of these, Davidson claims, require that the subject be able to attribute beliefs to himself. For understanding the possibility of error involves the recognition that one has beliefs that could be false; and surprise at being mistaken involves the recognition that certain of one’s beliefs have turned out to be false (and must therefore be given up).

Second, since it is crucial to the concept of belief that beliefs can be false, possession of that concept involves a ‘command of the subjective-objective contrast’ (Davidson 1982/1985, p.480), and this can develop only in the context of linguistic interpretation (Davidson 1975/1984, pp.168-70). In order to get a grip on what Davidson has in mind here, it is useful to take up Nozick’s important observation that the objective can reasonably be understood as what is ‘invariant under various transformations’ (Nozick 1998, p.21). According to Nozick this explains the ‘three strands to our ordinary notion of an objective fact’, which are as follows.

First, an objective fact is accessible from different angles[...]. It can be repeated by the same sense ... at different times, ... can be repeated by different senses of one observer, and by different observers. ... [Second,] there is or can be intersubjective agreement about it. ... [Third,] it holds independently of people’s beliefs, desires, hopes and observations... (Nozick 1998, p.21).

With this in the background it is reasonable to take Davidson as claiming, in effect, that our grasp of the concept of objectivity depends upon our ability to triangulate the objective through intersubjective agreement (cf. Davidson 1982/1985, p.480), and that this ability emerges only with our involvement in linguistic communication and interpretation (Davidson 1975/1984, pp.169-70 and 1997, p.27).

If the second stage of the argument is justified, then only linguistic beings can have the concept of belief. But the conclusion of the first stage is that any thinker, or being which has beliefs, must itself have the concept of belief. Taken together these imply that only linguistic beings are thinkers.

3. First-Pass Response
In order to undermine Davidson’s case for the thesis that thought depends upon language, it is not enough to show that each of the above arguments does not have compelling force (which is, of course, true). For even if each argument merely increased
the credibility of the thesis, their combination could add up to a significant case. However, as I will show in this section, it is easy enough to answer Davidson without requiring him to meet unreasonably high standards if one is open to the idea that some nonlinguistic animals have beliefs. This is more than enough to cast doubt on Davidson’s arguments as I have presented them above. The issue of whether it is possible to refashion those arguments into a stronger challenge to the idea that nonlinguistic animals could have beliefs, is one that I take up in later sections.

With respect to his first argument on the first front, Davidson’s assumption that we attribute beliefs to nonlinguistic animals on the basis of their observable behaviour alone is in fact false. We are also significantly constrained by at least two further factors. First, there is our knowledge of the information that is likely to be available to the animal concerned through perception (which is based on our own access to the environment, and what we know about its changing orientation, the sensitivities of its perceptual organs and its responses to various stimuli). This provides an excellent clue to some of its beliefs. Second, there is our knowledge about its standing and occurred biological needs and drives, which reveal many of its actual and possible desires. Given both these types of knowledge, our attributions of beliefs and desires to a nonlinguistic animal need not be subject to radical indeterminacy.

It is obvious, for example, that the explanation of the dog’s barking up the oak tree cannot be that he believes that the baby is there, and that if he barks at her for a while she will drop him a copy of the Critique of Pure Reason, which he really wants to read. For even though these beliefs and desires would (as Davidson puts it) ‘rationalise’ the dog’s behaviour, they are ruled out by a number of palpably nonlinguistic facts: that it was the cat, not the baby, that the dog pursued and saw running towards the tree; that dogs are not differentially responsive to copies of the Critique of Pure Reason; and that (given their biological natures) they cannot read or want to read. Of course this still leaves room for some indeterminacy with respect to the dog’s beliefs and desires. But there is no reason to think that this indeterminacy significantly exceeds that pertaining to everyday attributions of beliefs and desires to human beings. For the additional constraints that apply when one takes human linguistic behaviour into account are largely offset by the extreme plasticity of human belief and desire, and (with that) the potential for irrationality which gives human beings the status of rational animals.

To turn to Davidson’s second first-front argument, I claim that when we seriously attribute a belief to a non-linguistic animal, it is both possible and necessary to treat that belief as part of an extensive and developing system of beliefs and desires. If we say that the dog believes that the cat is up a certain tree when she is not in view there, we will also be ready to ascribe various related beliefs and desires to him. He wants to chase, catch, maol or bite the cat. He believes that the cat can climb, and that she is apt to reappear after disappearing. A moment ago he believed that the cat was running towards the tree. After a while, when confronted with new evidence, he will come to be-

---


12 It is also, incidentally, worth noting that some quite sophisticated intentional activities could in principle be expressed by complex nonlinguistic behaviour. It is possible, e.g., to describe hypothetical behaviour by a nonlinguistic animal that expresses the weighing of probabilities (see Jeffrey 1985, pp.484-5) and the systematic acquisition of counterfactual information (see Martin 1987, p.287; also Heil 1992, pp.223-5).
lieve that the cat is climbing into what we know as the kitchen window. He will then abandon the tree, lose his desire to pursue the cat, develop the desire to have a drink and lie down in the shade, and head off in the direction of the porch on the far side of the house, where he expects to find drinking water. In the background are various more-or-less standing beliefs about the layout of his home environment which allow him to rush around the corner without danger of bashing into a brick wall, or to move directly to the hidden gap in the fence when he wants to get out of the garden, and so on. If he has any beliefs at all, then he must have a developing system of beliefs and desires. Davidson is right about that. But he is wrong to think that we cannot reasonably attribute such an intentional system to a nonlinguistic animal — even though it will undoubtedly be a much less complex system than the intentional system of a normal human being.

In response to Davidson’s third first-front argument, I claim that semantic opacity is easily possible in the case of nonlinguistic beings. A dog could believe that the man approaching him with a stick is about to attack him without believing that his master is about to attack him, even though that man happens to be his master. Of course I do not mean to suggest that the phrases ‘the man approaching him with a stick’ and ‘his master’ perspicuously express the relevant canine mental contents, and I use them only to distinguish between the dog’s having a belief about someone via an occurrent perception and its doing so via the mechanism (whatever it may be) through which it normally recognises the man whom we describe as its master. More generally, it ought to go without saying that attributions of beliefs to nonlinguistic animals cannot hope to express those beliefs, but only to describe them roughly, and from our own point of view (cf. Dennett 1996, pp.41-3 and Dretske 1988, pp.121, 155). This recognition clearly muddies the application of the substitution test for semantic opacity in the case of nonlinguistic animals, but this does not invalidate the above example. What it does, rather, is point to the limitations of the substitution test, and the need for an alternative account of semantic opacity (which I provide in section 5).

On Davidson’s second front I would like to consider the second stage of his argument first. Heil shows convincingly that it is not a conceptual truth that an appreciation of the subjective-objective contrast, and therefore the concept of belief, can develop only in the context of linguistic communication and interpretation (Heil 1992, pp.215-20). I will not repeat his arguments, but the idea of a solitary intelligence which appreciates the contrast and possesses the concept is surely not incoherent, even though such a being is completely implausible. In the light of his rejection of the analytic-synthetic and scheme-content distinctions (see Davidson 1974/1984) it is, however, doubtful that Davidson meant to be advancing a purely conceptual thesis. I suspect that he meant to be talking about a fairly restricted but completely unspecified class of possible beings that does not extend much beyond animals roughly like those we know about or, perhaps, biologically possible animals, and that he meant to claim on broad empirical grounds that any such being that develops an appreciation of the subjective-objective contrast does so in the context of linguistic communication.

This claim is in line with Dreckmann’s (1999) careful argument that the behavioural evidence does not warrant the attribution of higher-order intentionality to nonlinguistic animals. The claim is, moreover, plausible in its own right. For it is hard to imagine how a normal child could begin to acquire the concept of belief without hearing speak-

13 However, this does not rule out the possibility of nonlinguistic animals having other forms of self-consciousness. See, e.g., Brittan 1999, pp.72-4.
ers making belief attributions in the medium of a public language, sometimes on the strength of assertions by the relevant subject. It is even less clear how she could develop an appreciation of the indispensable idea that beliefs are (for the most part) subject to evaluation for truth or falsity on the basis of evidence without witnessing speakers’ involvement in the communicative practices constituting the social dimension of ‘the space of reasons’ (Sellars 1963, p.169; see also McDowell 1994, pp.4-5). This includes, in particular, their advancing reasons for or against the beliefs that they attribute to others, and their evaluating, criticising and correcting each others’ reasons and reasoning.14

There is, therefore, considerable merit to Davidson’s views on the conditions necessary to develop an appreciation of the subjective-objective contrast, which is inseparable from the concept of belief. However, the first stage of his second-front argument is much less secure. Beliefs are clearly subject to truth and error, but Davidson has given us no good reason to suppose that every believer must appreciate this. We have no great difficulty in attributing both true and false beliefs to animals like dogs, which surely lack the idea of truth – and the ideal of objectivity that marks the capacity for judgment. Moreover, we can attribute both true and false beliefs to such animals even under the reasonable supposition (which may have been in the back of Davidson’s mind) that candidate beliefs cannot be genuine unless their truth or falsity is of consequence to the believer. For this condition does in fact apply if the believer’s success at satisfying its needs and wants depends in part on the extent to which its beliefs are true, and its sensory and cognitive mechanisms and processes tend to produce beliefs that are predominantly true. An explicit desire for truth and an idea of objectivity are not also required.

Furthermore, it makes no difference that believers must, as Davidson rightly claims, be capable of surprise when they are mistaken. The dog, having believed that the cat is up the tree, can be surprised to see her jumping into the kitchen window without being capable of thinking that his belief was false. For nothing is required to make sense of the dog’s surprise other than a cognitive tension between his perception of the cat’s jumping into the window and the residue of his disposition to believe that the cat is in the tree.

4. Heil’s Davidsonian Reply

Although responses to Davidson like the above are necessary, they fail to engage with a deeper and more interesting challenge underlying Davidson’s thinking which has been brought to the surface by Heil (1992, especially pp.198-201, 214-5).

The challenge arises from the fact that many different kinds of things have ‘representational’ states in the sense that they carry information (and misinformation) that determines or influences their behaviour. This includes thermostats, trees, autonomic nervous systems at all levels of complexity, slugs, snakes, dogs, chimpanzees, and human beings from infancy to senility. And some of their representational states are thoughts, including Maropeng’s belief that Sam is in the kitchen, and Davidson’s judg-

---

14 These reflections are, I believe, consistent with available evidence in cognitive developmental psychology. They are not, however, positively confirmed by that evidence, which is mainly about how children use and understand certain key words and expressions at various ages, without systematic information concerning conditions and circumstances that may have influenced the development of the relevant semantic competences. See Woodfield 1996 for a brief overview of some of the empirical evidence along with useful philosophical discussion of how it is best interpreted.
ment that thought depends upon language. But others are not, like the thermostat's registering that the temperature in the refrigerator is above 6°C, the nervous system’s signal of food in the stomach, and the slug’s representation of greater moisture content in the soil above. Again, some representers can think, like most normal human beings, while other cannot, like trees, slugs and autonomic nervous systems.

The problem is to give a principled account of what it takes for a representation to qualify as a thought, and a representor as a thinker. This is a significant challenge for contemporary philosophy of mind even aside from the question of whether thought depends upon language. And it is a challenge that is especially pressing for those who accept the dominant view that intentional states are functional states, which is significantly driven by the recognition of how easy it is to provide functionalist accounts of much simpler representations.

Davidson’s case for the thesis that thought depends upon language contains the germs of an account of what distinguishes thinkers from mere representers, albeit one that we may consider too demanding. The limitation of my section 3 response to Davidson is that it contributes very little towards the development of a reasonable alternative. Before I turn to the question of how this gap should be filled, I want to look briefly at Heil’s attempt to reconstruct a defensible Davidsonian account.

Heil takes his cue from Davidson’s view that thinkers must have the concept of belief and be capable of self-consciousness (see Davidson 1982/1985, p.478):

The having of thoughts is not merely a matter of representing ... but of appreciating in some way that this is what one is doing, of exhibiting the capacity to manipulate and act on representations qua representations. An appreciation of this sort is involved, for instance, in one’s comprehending a distinction between representations and represented states of affairs, between one’s subjective impression of the world and the world as it is independently of that impression, between opinion and truth (Heil 1992, p.200).

Now it is easy to concede that any being that can appreciate that it has representations has thoughts, and that this standard clearly separates us from thermostats, slugs and autonomic nervous systems. But Heil’s suggestion is nonetheless problematic. For, although appreciation that one is a representor is clearly sufficient for thought, anyone who is open to the idea that nonhuman animals could have thoughts may reasonably doubt that it is also necessary. Furthermore, since appreciating trivially involves thinking, the claim that beings that appreciate that they have representations are thinkers is actually uninformative.

15 As Dretske (1988, pp.64-70) insists, anything capable of representation is also capable of misrepresentation. Thus the mere possibility of error is not sufficient to distinguish beliefs from other representations.
16 Consider, e.g., Heil’s (1992, pp.198-205) account of ‘weak representation’ and Dretske’s (1988, Chapter 3) account of simple representational systems. Perhaps the most impressive attempt from a functionalist perspective to address what I have described as the problem of distinguishing thinkers from other representors appears in Dretske 1988, Chapters 4-6, and I draw on Dretske’s insights in section 6. For a brief outline of some of his leading ideas, see Dretske 1999.
17 Heil’s aim is ‘to make Davidson as plausible as possible — most particularly Davidson’s ... contention that belief requires the concept of belief’ (personal communication). He is not committed to the truth of his reconstructed Davidsonian account, and he firmly rejects the thesis that thought depends upon language (see Heil 1992, pp.223-5).
No doubt Heil appreciates these points, for the quoted passage serves merely as a first guiding step in a careful discussion of the issues. This is not the place to work through the subtleties of Heil’s thinking, but, roughly speaking, the position he arrives at replaces Davidson’s claim that only beings with the concept of belief have beliefs with the thesis that only representational systems with a developed capacity for genuine second-order representations of their own first-order representations have thoughts.

Heil attempts to underwrite this by the remarkable idea that the semantic opacity of thought may be rooted in second-order representation. Explicitly ‘equat[ing] semantic opacity with fine grainedness or definiteness of content’ (Heil 1992, pp.194-5), Heil reasons as follows:

A genuine second-order representation includes in its content first-order representational content. In this respect a second-order representation settles the issue of definiteness. A representational state that failed to do so would not be a second-order state, a representation of a representation. A system that incorporates second-order representations, then, necessarily possesses fine-grained intentional capacities... (Heil 1992, p.204).

The suggestion that the semantic opacity of thought is due to second-order representation is interesting and inventive. It also yields an extremely neat unification of the two fronts of Davidson’s campaign for the thesis that thought depends upon language by positing a constitutive link between holism and opacity on the first front and self-consciousness on the second.

5. Problems with Heil’s Proposal

Heil’s idea is attractive, but I do not think it works. More specifically, there is good reason to believe that second-order representation is neither necessary nor sufficient for semantic opacity.

In order to engage with the issues we need an account of semantic opacity that does not (like the substitution test) presuppose that the representations concerned can be expressed in a familiar language like English. What I would like to suggest is that in the basic case a representational system is opaque if it includes two or more distinct ways of representing one and the same state of affairs. Generalising from this idea, we may define a representational system as semantically opaque if and only if it involves at least two possible representational states, R and S, with the following two properties. First, R and S are independent in the sense that the representer could be in either of these states without being in the other. Second, R and S are equivalent in the sense that they are satisfied by the same possible state of affairs. I shall call this the direct account as it defines semantic opacity in terms of properties of the relevant representations rather than via properties of attributions of those representations.

Because it utilises the idea of satisfaction rather than mere truth (which is but one species of satisfaction), the direct account applies not only to what Millikan calls ‘descriptive’ representations, like beliefs – which ‘represent ... what is the case’ – but

18 I argue for this more carefully in Pendlebury 2002, and section 5 of the present article is a summary of the key thoughts of that paper, to which the reader should refer for further details and important qualifications. I also draw on Pendlebury 2002 with respect to some necessary background to section 5 that appears in earlier sections.
also to ‘directive’ representations, like desires – which ‘represent ... what is to be done’ (Millikan 1996, p.145). This is as it should be.

Any form of representation that is opaque in terms of the substitution test will also be opaque in terms of the direct account. To illustrate this in the case of belief using my earlier example: the belief that the first Postmaster General of the United States was a statesman and the belief that the inventor of bifocals was a statesman are independent because a subject could have either of these beliefs without having the other; but the beliefs are equivalent because they are satisfied under the same conditions.

Furthermore, the direct account of semantic opacity can be applied even when the substitution test cannot, and in undisputed cases it yields appropriate results. For example, a regular thermostat is not capable of a broad enough range of representational states to meet the specified conditions of semantic opacity. And without presupposing that animals like dogs have beliefs, it is possible to use the account to make sense of the idea that they could have semantically opaque representational systems (cf. section 3 above). Thus, all things considered, the direct account of semantic opacity is clearly superior to the substitution test.

On the basis of the direct account, it is reasonable to hold that second-order representation is not necessary for semantic opacity. Some might want to invoke animals like dogs in support of this position on the ground that they have semantically opaque representational systems without having second-order representations, but at this point of the debate that claim is too contentious to depend on. It is, however, easy to specify an imaginary representational system that is semantically opaque without being capable of higher-order representations, and thus to avoid disputes induced by conflicting intuitions about familiar animals and artifacts.

Consider, then, the case of a computer-based representational system which automatically assembles detailed records of the financial affairs of many individuals by accessing various computer networks, and which automatically responds to loan applications entered by the individuals concerned on bank terminals by granting the loan applied for, refusing the application out of hand, or referring the applicant to a loan officer in a bank that participates in the system – and doing this systematically on the basis of the information in its records. It is easy to imagine that such a system might not be able to rule out the possibility of, so to speak, different individual ‘markers’ or ‘files’ which apply to one and the same individual. If so, then the system is semantically opaque, for it has independent ways of representing the same state of affairs; and this holds regardless of whether it is capable of second-order representations.

It is also reasonable to hold that second-order representation is not sufficient for semantic opacity. Consider a sophisticated temperature control system for a hotel room which switches the heating and airconditioning on and off at temperatures selected by the guest, but which also incorporates self-monitoring devices, the outputs of which

---

19 Millikan’s distinction between descriptive and directive representations corresponds to Searle’s distinction between representations with a mind-to-world direction of fit and those with a world-to-mind direction of fit (see Searle 1983, p.8). Although the terminology of descriptive and directive representations provides a convenient way to mark an important distinction, it is generally inappropriate if taken literally. Beliefs, e.g., are not descriptions, and desires are not directions.

20 My appeals to imaginary representational systems in this section also help to ensure that Heil’s very reasonable constraints on representation (Heil 1992, pp.198-9) are satisfied. I discuss the issues in Pendlebury 2002.
automatically have various results, such as overriding the guest’s instructions in favour of predetermined fallback settings under certain conditions, shutting down the whole system under others, and charging the guest’s account on the basis of the duration of the periods during which the heater and air conditioner are activated as well as the temperatures at which they are activated (so that fuel economy can be promoted by a double incentive). It is true that the contents of some of the self-monitoring devices in such a system could reasonably be understood as being about the environment, the heater or the air conditioner rather than about the system’s own first-order representational states. But this does not apply to all the candidate second-order representational states in the system. Consider in particular self-monitoring states such as that which (as it were) says that the system is in the directive first-order representational state that switches the air conditioner on when the ambient air temperature is 25°C or higher. Given their potential impact on the system’s override and shut-down functions as well as on the billing system, such self-monitoring states must be treated as second-order descriptive representational states, the contents of which are about the system’s first-order directive representational states. But the system is not semantically opaque in terms of the direct account, since it is not capable of independent representational states that are satisfied by the same states of affairs.

The reasoning of this section not only supports the view that second-order representation is neither necessary nor sufficient for semantic opacity, but also shows that neither semantic opacity nor second-order representation is sufficient for thought (for it is obvious that neither of my hypothetical representational systems can think). In addition, it suggests that semantic opacity is a matter of degree (depending on the range of equivalent alternative representations that are possible in the system), and that the mere presence of opacity is no guarantee that the system is capable of the fine-grained distinctions exhibited by natural languages and distinctively human forms of thought.

The question of what degree of semantic opacity is required for thought is left open by the above considerations, and it may be doubted that on its own any degree of semantic opacity is sufficient for thought. Furthermore, none of the reasoning that I have considered supports the view that second-order self-consciousness is required for all thought – although it is no doubt inseparable from the high-level thought involved in human reasoning and judgment.

6. A Better Alternative

Having cast doubt on the suggestion that second-order self-consciousness is necessary for a representor to qualify as a being that thinks, I owe the reader some ideas on a possible alternative account. If it is to underwrite the arguments advanced in section 3 in response to Davidson, this alternative must be one in terms of which some nonlinguistic beings can think.

I regret that I do not have a neat, unifying theory to offer. However, such a theory may be impossible, especially given the vagueness of the word ‘thought’, which is obviously not a natural-kind term. Thus, instead of attempting to adumbrate a well-integrated theory of thought, I would like to present an overview of some conditions that seem to me necessary for a representor to qualify as a thinker, and then to reflect briefly on what conditions should be counted as clearly sufficient. Some of the
considerations that I will mention have been foreshadowed in earlier sections of this article, and some draw on important recent work on cognitive ethology.\textsuperscript{21}

I will assume that a reasonable conception of a thinking being should be broad enough to cover animals (including higher mammals) to which we often, freely and seriously ascribe intentional states like beliefs and desires in everyday speech. But it should also be narrow enough to make good sense of our disposition to classify them together, and to do other useful work. It should, for example, offer significant prospects of interesting and worthwhile patterns of explanation and prediction which apply more or less generally to the items it classifies together as thinkers,\textsuperscript{22} regardless of differences between them, but which do not apply to items that it clearly excludes from this class. These methodological considerations suggest that we should deny something the status of a thinker in the absence of good reasons to count it as one.

Among the most obvious necessary conditions for a representer to count as a thinker are that it should display a wide range of complex and variable behaviour – that we can explain and predict only via the intentional stance – and that the relevant representational system should be extensive, complex, semantically opaque, and subject to ongoing change. As suggested earlier, it must have needs and drives that are revealed by biological considerations, features of its environment and its physiological states, and its candidate desires must accord with those needs and drives.\textsuperscript{23} It must also be a perceiver, and its candidate beliefs must be consonant with and to a significant degree explicable on the basis of its past and present perceptions.

More particularly, to qualify as a thinker it is crucial that a representer should not only be responsive to immediate perceptual information, but also that it should harbour particular and general information about its environment that is not available to it perceptually, but is manifest in its behaviour, explicable in terms of its perceptual history, and subject to change in response to new perceptions. It must, therefore, display ongoing learning that goes beyond simple stimulus conditioning like that exhibited by molluscs of the species \textit{Aplysia}, whose defense reflexes (involving the withdrawal of their respiratory siphons into the mantle cavity) can be modified by habituation and sensitisation (see Proust 1999, pp.42-4). More particularly, it must display a capacity for instrumental learning through natural or artificial reinforcement (as when a rat learns that it can get food by pressing a bar as a result of its getting food unexpectedly on the first few occasions when it presses that bar) as well as a capacity for tacit learning by observation without reinforcement (as when a rat learns its way through a maze without rewards and punishments) (see Dretske 1988, pp.115-9, 144-5).

Furthermore, it is crucial that the ongoing changes in the representer's candidate belief system should reveal a robust \textit{objective pull}. Part of what I mean by this is that the

\textsuperscript{21} \textit{Erkenntnis} 51, No.1 (1999) contains a good selection of recent papers at the interface between cognitive ethology and philosophy of mind, including several papers cited in this article. Dretske 1988 draws extensively on earlier work in the field.

\textsuperscript{22} I use the qualified form 'more or less generally' because the class of thinkers will not be sharply defined, is not a natural kind, and may in part be constituted by family resemblances.

\textsuperscript{23} This reference to biological considerations and physiological states is not meant to exclude the possibility of an artifact being a thinker, but only to imply that in order to qualify as one, an artifact (in addition to satisfying all other necessary conditions) must in its own right exhibit homeostasis and goal-directedness of a sufficiently robust form for it to have needs and drives. I am skeptical of the chance of any such artifact being developed by non-genetic means in the short to medium-term future.
operation of the underlying processes and mechanisms should be apt to increase the probability of truth in the system, especially by correcting misinformation in response to perceptual evidence – as when the dog stops believing that the cat is up the tree on seeing her jump into the kitchen window. The capacity for error detection and correction can be quite sophisticated in nonlinguistic animals. Allen mentions the case of pigs being trained to ‘respond differentially to stimuli consisting of a pair of objects depending on whether the two objects differed in … size, shape or colour, or whether they were identical in those respects’ (Allen 1999, p.37):

[These pigs] would still occasionally make errors (wrong choices) while performing at an overall rate at or near 90% correct. After committing to a response but before any feedback was provided, some pigs would attempt to back away from the choice they had made. Analysis of 22 cases of backout behaviour showed that only one of these cases occurred after the pig made a correct choice... (Allen 1999, p.38).

This provides evidence that the pigs concerned were capable of a rudimentary form of spontaneous ‘self-monitoring’ which falls short of second-order self-consciousness, but such a capacity is not essential for an animal to be responsive to perceptual evidence.

A deeper, more Kantian feature of the objective pull of belief is connected with the fact that genuine belief systems are concerned not only with proximal information, but even more notably with distal information.24 Such information is crucial for the explanation of behaviour involving the (largely successful) orientation of believers towards a world of ‘stable objects and changing properties’ (Proust 1999, p.45) in objective space and time. Distal information must be subject to some form of triangulation by the believer. Or, to paraphrase Nozick (see section 2), it must be invariant under certain transformations, and be accessible via the same sense at different times and different senses at the same time. In the actual world such access depends upon a process of ‘calibration’ of different senses against each other that ‘allows the animal to correct its sensory inputs and puts them into one single spatially coherent perceptual frame.’ (Proust 1999, p.51) There is empirical evidence of general calibration mechanisms in birds, reptiles and mammals. Proust mentions the example of

how a young owl whose correspondence between visual and auditory stimulus representation (in the tectum) has been experimentally distorted (through a plug in one ear) will restore spatial congruence: ‘the projection of the auditory stimulus onto the tectum gradually changes so as to bring the effect of a sound back in register with the effect of a visual stimulus originating at the same angular deviation’ (Gallistel 1990, p.481) (Proust 1999, p.51).

It may be that even the most primitive capacity for error correction depends upon an objective spatio-temporal orientation made possible by calibration (or something analogous to such an orientation),25 for it is unclear that beings that are responsive only to

---

24 Here I draw on ideas advanced in an interesting and suggestive paper by Proust (1999), but I adapt them to my own purposes and clothe them to some extent in my own terminology.

25 This parenthetical qualification is a concession to Strawson’s (1959, Chapter 2) recognition that something other than space could play the same role as space in securing objectivity. It may, however, be doubted that objective experience is possible in the purely auditory world imagined by Strawson precisely because only one sense is involved (see Proust 1999, p.47).
proximal information have ‘representations about the world that could be mistaken.’ (Proust 1999, p.43).

The objective pull of belief is but one side of the fact that (as I put it in section 3) the truth or falsity of beliefs are of consequence to the believer, the other being that the believer’s success at satisfying its wants and needs depends significantly on the extent to which its beliefs are true. I would now like to suggest that the idea that genuine beliefs must be of consequence to the believer is best understood as a particular case of the broader idea that the satisfaction or non-satisfaction of genuine intentional states must in general be of consequence — or make a difference — to the subject. This would go a long way towards distinguishing thinkers from other representers. For although it makes no difference to a thermostat, a present-day computer or a cat’s autonomic nervous system whether its descriptive states are correct or its directive states are effective, it certainly makes a difference to a human being, a chimpanzee or a cat whether its beliefs are true and its desires fulfilled. And it does so even if the individual does not, or cannot, have higher-order intentional states concerning its first-order representations.

But in virtue of what is desire-satisfaction of consequence to the subject while the satisfaction of non-intentional directive states is not? The answer cannot be that desires aim, as it were, at their own satisfaction, since this is true of all directive states. Nor is it that desire-satisfaction contributes directly to the biological health of the deseriper, even though this would rightly exclude the directive states of thermostats, present-day computers and autonomic nervous systems from the class of desires. For one thing, desire-satisfaction can sometimes be bad for one’s health. For another, plants have directive representational states that are not desires even though their satisfaction contributes to the health of the plants. Thus for our purposes it would not do to regard the satisfaction of these states as making a difference, or being of consequence, to the plants.

What I would like to suggest is that the distinctive feature of desire in virtue of which its satisfaction makes a difference to the subject is that it is appropriately connected with feelings of pleasure and pain (understood very broadly to include feelings of comfort, discomfort, agony, delight, irritation, relief and the like). More particularly, directive representational states are of consequence to the subject and so qualify as desires only if the subject is capable of pleasure and pain, is drawn to the former and discouraged by the prospect of the latter, and is generally apt to experience a more favourable balance of pleasure over pain when those directive states are satisfied, and a less favourable balance when they are not. This is in line with the fact that in our everyday thought and talk, desire is pervasively connected with feelings of pleasure and pain. There is, therefore, nothing farfetched in the idea that a necessary condition for a being to have genuine desires is that it be subject to such feelings. If this is correct, then since belief is impossible without desire, it follows that pleasure and pain are necessary for thought.

This reasoning makes sense, however, only if pleasure and pain are understood as affects that are felt by the subject rather than as functional states that are exhausted by their normal causes, effects and explanatory roles. Whatever their physiological basis, in the present context we must view them as having an ineliminable phenomenal, qualitative and subjective dimension. In the language of current philosophy of mind, they necessarily involve qualia. I am, therefore, proposing that in order for a representor to
qualify as thinker, or subject of intentionality, it must have qualia, including pleasures and pains.

This proposal is at odds with a tendency in philosophy of mind towards segregating issues of intentionality from supposedly residual issues of ‘phenomenology’ concerning qualia alone. It is as if there is a background presumption that we can or ought to be able to make sense of intentionality without reference to qualia, which presents special problems of accommodation; and that the burden of proof automatically falls on those who disagree. This shows up in a variety of ways. Davidson, for example, has nothing to say about non-intentional aspects of mind, Dennett (1991, Chapter 12) denies the existence of qualia, and Dretske (1995) and Chalmers (1996) engage in extended efforts to find a place for them in nature along with intentional states, which are assumed to be much more securely at home there. The segregation of the intentional and the phenomenological is ill-founded if I am right in thinking that genuine thinkers must have experiences of pleasure and pain.

There are evidently significant differences between representers that clearly satisfy all the necessary conditions for thinkers proposed above and those that fail to satisfy many of them. Furthermore, the former class includes all the beings that most of us are strongly disposed to count as thinkers, namely, human beings and higher mammals, while the latter includes all representers to which we are naturally inclined to deny that honorific, including plants, very primitive animals, autonomous nervous systems, and the artifacts we know about. This suggests not only that my proposed necessary conditions are on the right lines, but also that as a tentative working hypothesis it is reasonable to adopt the view that compelling satisfaction of all of these conditions is enough for a representor to count as a thinker.

The considerations sketched in this section show how it is possible to answer the challenge to distinguish between thinkers and other representers without supposing that all thinkers must be subjects of second-order self-consciousness or speakers of a language. They also provide good reason to reaffirm the view that some nonlinguistic animals can think.

7. Coda
I would like, finally, to suggest that language itself presupposes nonlinguistic thought, and therefore that reflective human thought, which is inseparable from language, is impossible in the absence of nonlinguistic thought.26

It is crucial that public natural language – the form of language with which we are concerned – involves norms or conventions that cannot be explained in terms of explicit rules formulated in language on pain of circularity. This commitment to linguistic norms is consistent with the current trend in linguistics, led by Chomskian minimalism, towards viewing language as an almost wholly non-normative, natural product of a biological faculty of language.27 For even if it is true that the fundamental properties of any natural language are determined largely by universal features of this faculty, and that ‘A rational Martian scientist would probably find the variation [between earthly languages] rather superficial, concluding that there is one human language with minor variants’ (Chomsky 2000, p.118), it is undeniable that there are dif-

26 See Pendlebury 1998 for a detailed presentation of the following schematic argument for this claim.
27 See Chapter 1 of Chomsky 2000 for a useful overview of these trends.
ferences between different languages that are a matter of convention. For instance, the distinctive ways in which a language expresses the concept of a dog, the first person singular, causation, pluralisation, negation, universality, or grammatical subjecthood are all manifestations of norms of that language that competent speakers have mastered.

Now all norms presuppose the existence of norms that are not reducible to explicit rules in public languages, for the application of any rule is itself subject to assessment as correct or incorrect on the basis of norms that go beyond the rule itself. It is reasonable to hold that such implicit, pre-linguistic norms are constituted by social practices involving not only the relevant behaviour that could be judged appropriate or inappropriate in terms of those norms, but also in assessments of that behaviour and associated reinforcement activity.\(^{28}\)

It is, however, unimaginable that we could make sense of such complex patterns of activity without assuming that the participants are capable of purposive activity, that they have beliefs and desires, and that they are subject to pleasure and pain. They must therefore be thinking beings, and the relevant kind of thinking cannot hinge on their having a prior mastery of the norms of a public language.\(^{29}\) Thus any thought that depends upon language presupposes nonlinguistic thought.\(^{30}\)

References


\(^{28}\) For a rich and subtle development and defense of this position, see Brandom 1994, pp.18-64. The main lines of Brandom’s arguments are summarised in Pendlbury 1998, pp.143-5.

\(^{29}\) This does not of course rule out the possibility that the kind of thought involved in the social practices in which human norms are rooted somehow depends upon the faculty of language itself.

\(^{30}\) I have presented material contained in this article to the Annual Conference of the Philosophical Society of Southern Africa, Durban, January 2001, the Spring Colloquium in Philosophy, Rhodes University, Grahamstown, September 2001, and the Department of Philosophy, University of the Witwatersrand, Johannesburg, October 2001, and am grateful to my audiences for useful questions and comments. I owe special thanks to Tom Stoneham for saving me from a couple of errors.


