

# *Radical interpretation, scepticism, and the possibility of shared error*

**Joshua Rowan Thorpe**

## **Synthese**

An International Journal for  
Epistemology, Methodology and  
Philosophy of Science

ISSN 0039-7857

Volume 196

Number 8

Synthese (2019) 196:3355-3368

DOI 10.1007/s11229-017-1600-z



**Your article is protected by copyright and all rights are held exclusively by Springer Nature B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Radical interpretation, scepticism, and the possibility of shared error

Joshua Rowan Thorpe<sup>1</sup>

Received: 9 May 2017 / Accepted: 21 October 2017 / Published online: 28 October 2017  
© Springer Science+Business Media B.V. 2017

**Abstract** Davidson argues that his version of interpretivism entails that sceptical scenarios are impossible, thus offering a response to any sceptical argument that depends upon the possibility of sceptical scenarios. It has been objected that Davidson's interpretivism does not entail the impossibility of sceptical scenarios due to the possibility that interpreter and speaker are in a shared state of massive error, and so this response to scepticism fails. In this paper I show that the objection from the possibility of shared error rests on a misunderstanding of Davidson's interpretivist position. Properly understood, Davidson's view does entail that sceptical scenarios are impossible. I also give a reason independent of its anti-sceptical implications to prefer Davidson's interpretivism over the version of interpretivism erroneously attributed to him (at least implicitly) by those who object to his anti-sceptical argument.

**Keywords** Epistemology · Scepticism · Donald Davidson · Philosophy of mind · Interpretivism

Davidson held that a central aspect of what it is for a subject to produce a sentence with a particular meaning is that a fully informed interpreter would conclude that the sentence has that meaning. Likewise, he held that it is a central aspect of what it is for a subject to have a particular belief that a fully informed interpreter would conclude that the subject does indeed have that belief. One consequence of this account of belief and meaning is that 'what a fully informed interpreter could learn about what a speaker means is all there is to learn; the same goes for what the speaker believes'.<sup>1</sup> That is, a sentence in a speaker's language has a particular meaning if and only if a

---

<sup>1</sup> Davidson (1986, p. 315).

---

✉ Joshua Rowan Thorpe  
[joshthop@gmail.com](mailto:joshthop@gmail.com)

<sup>1</sup> University of Campinas (Universidade Estadual de Campinas), São Paulo, Brazil

fully informed interpreter would conclude that it has that meaning, and a speaker has a particular belief if and only if a fully informed interpreter would conclude that she has that belief.<sup>2</sup> Let us call any view of content that has this consequence an interpretivist view.<sup>3</sup> Davidson characterised his version of interpretivism by specifying the method of interpretation that must be used by a radical interpreter, an interpreter who starts out with no knowledge of any semantic or intentional facts about the speaker who is to be interpreted. For Davidson, it is the conclusions reached by a radical interpreter that give a complete specification of the meanings of a speaker's sentences, and of a speaker's beliefs.

My concern in this paper is not to defend Davidson's interpretivism. Rather, I am interested in Davidson's claim that his interpretivist view offers a response to scepticism. Many arguments for scepticism rely, explicitly or implicitly, on the claim that sceptical scenarios in which a thinker has all (or at least mostly) false beliefs about her environment are possible. The case of a brain-in-a-vat fed sensory experiences by a supercomputer is usually taken to be an example of such a scenario. Davidson (1986, p. 316) argues since the radical interpreter must 'read some of his own standards of truth' into the utterances and beliefs of the speaker, sceptical scenarios are impossible.<sup>4</sup> If this entailment holds, and if Davidson's interpretivism is true, this would have the significant consequence that all sceptical arguments which rely on the claim that it is a logical possibility that I am in a sceptical scenario are unsound.

This would be a powerful anti-sceptical result, and one might think that the promise of such a result would encourage epistemologists to take an interest in the question of whether Davidson's interpretivism *is* true. However, the Davidsonian response is largely neglected in contemporary epistemology, due mainly to the fact that it is widely thought that the crucial entailment does not hold. It has been suggested that even if we accept Davidson's interpretivism his argument that sceptical scenarios are impossible fails, due to the possibility that both the speaker and the radical interpreter are in a shared state of massive error.<sup>5</sup> Davidson's remarks about an omniscient interpreter are usually perceived to be an attempt to deal with this problem, but this perceived attempt is widely thought to be a failure.<sup>6</sup>

The primary purpose of this paper is to show that the crucial entailment does hold. The objection from the possibility of shared error depends upon the attribution to

<sup>2</sup> As Byrne (1998) notes, it is possible to read 'all there is to learn' in this quotation in a way that leaves open the possibility that a speaker has beliefs and meanings that cannot be discovered by an interpreter. On this reading, the possibility of the facts about a speaker's meanings and beliefs outrunning what a fully informed interpreter could learn is compatible with Davidson's claim. However, Davidson clearly did not intend his claim to be read in this way. By 'all there is to learn' he simply meant all the facts.

<sup>3</sup> Other interpretivists include Lewis (1974), Dennett (1971) and Childe (1994). It is an interesting question whether these interpretivist views might form the basis of a response to scepticism. However, in this paper I will focus solely on the anti-sceptical implications of Davidson's interpretivism.

<sup>4</sup> The most explicit statement of this argument is in Davidson (1986), and I will focus on the version of the argument given in this paper. However, versions of the argument also appear in Davidson (1984, 1991a, b, 2005).

<sup>5</sup> Stroud (2000, 2011), Ludwig (1992), Brueckner (2010) and Williams (1988).

<sup>6</sup> Stroud (2000), Vermazen (1983), Foley and Fumerton (1985), Ludwig and Lepore (2005) and Williams (1988).

Davidson of a view that I will call agreement interpretivism, which requires only that the interpreter and the speaker agree, thus making it possible for someone with mostly false beliefs about a speaker's environment to apply the method of radical interpretation to that speaker correctly. However, Davidson did not endorse agreement interpretivism. He endorsed a view that I will call truth interpretivism, which requires that the interpreter attribute beliefs to the speaker that are actually true, and so makes it impossible for someone with mostly false beliefs about a speaker's environment to apply the method of radical interpretation to that speaker correctly. Thus, there is no need for a further argument from Davidson to rule out the possibility of shared error, and his remarks about an omniscient interpreter are not an attempt to provide such an argument. I conclude that, properly understood, Davidson's interpretivist view does provide a response to any form of scepticism that relies on the claim that sceptical scenarios are possible.

We shall also see that there is reason independent of its anti-sceptical implications to prefer truth interpretivism over the agreement interpretivist view erroneously attributed to Davidson (at least implicitly) by those who make the objection from the possibility of shared error. Agreement interpretivism seems to entail contradictions in a way that truth interpretivism does not. So although I will not attempt to vindicate truth interpretivism, I do draw the more cautious conclusion that it is preferable to agreement interpretivism.

## 1 The objection from the possibility of shared error

In this section, I will first describe the usual way of understanding the Davidsonian response to scepticism and the interpretivist view on which it depends. Some of this understanding is correct, but I will flag aspects of it that are incorrect or incomplete to which I shall return in the next section. I will then describe the objection from the possibility of shared error.

The Davidsonian response to scepticism is premised on his version of interpretivism, according to which a sentence in a speaker's language has a particular meaning if and only if a radical interpreter would conclude that it has that meaning, and a speaker has a particular belief if and only if a radical interpreter would conclude that she has that belief. Ultimately, the radical interpreter hopes to arrive at an interpretation taking the form of a Tarski-style truth theory that entails, for any assertoric sentence  $s$  in the interpreted speaker's language, a theorem of the form

(T)  $s$  is true iff  $p$

where  $p$  is replaced by a sentence in the language of the interpreter. Since an infinite number of well-formed and meaningful assertoric sentences can be constructed in a natural language, and since what can be learned by an interpreter must be finite, these theorems must follow from a finite number of axioms that specify the referents of proper names, the satisfaction conditions of predicates, and the rules for the logical vocabulary.<sup>7</sup>

---

<sup>7</sup> This specification of what sort of theory the radical interpreter wants to confirm is sufficient for our purposes. However, it might be thought that the theory must fulfil further conditions; see Ludwig and Lepore (2005, pp. 151–173) for discussion.

How might a radical interpreter go about confirming such a theory, given that she must start the process of interpretation with no knowledge of any semantic or intentional facts about the speaker? The answer that Davidson is usually thought to give to this question is that the interpreter must pay attention to the speaker's patterns of assent and dissent to sentences in her language. In particular, the interpreter will notice that the obtaining of certain conditions in the speaker's environment will cause the speaker to assent to certain sentences in her language. For example, the interpreter might notice that, usually, the presence of a cat causes a speaker to assent to 'lá se vai um gato'. These are the occasion sentences of the speaker's language. In Sect. 2 I will suggest that this account of Davidson's answer is incomplete. Patterns of assent to occasion sentences do play an important role in the process of radical interpretation. However, they are a sub-set of the total evidence available to the radical interpreter.

If the interpreter accepts certain guiding principles, occasion sentences will provide her with a starting point for interpreting the speaker's language. The first principle is that a speaker assents to sentences in her language, and to occasion sentences in particular, because, on the whole, those sentences express beliefs that she holds.<sup>8</sup>

Davidson (1986, p. 316) puts the second principle of interpretation by saying that the interpreter must read 'some of his own standards of truth into the pattern of sentences held true by the speaker.' More precisely, the interpreter identifies the truth conditions of the speaker's occasion sentences with the states of affairs that typically cause the speaker to assent to them.<sup>9</sup> Thus, the interpreter will see the speaker's occasion sentences as being, on the whole, true at the times when the speaker assents to them.

The first and second principles together allow the process of interpretation to get started with some provisional attributions of belief and meaning to the speaker. For example, upon noticing that the presence of a cat usually causes a speaker to assent to 'lá se vai um gato', the interpreter provisionally concludes by the first principle that this occasion sentence expresses one of the speaker's beliefs, and by the second principle that the truth condition of this sentence, and thus of the belief that it expresses, is that there is a cat around.

Once the interpreter has a fair number of conclusions of this sort in hand he will be able to begin to infer axioms governing individual terms—that is, the proper names, predicates, and logical vocabulary—of the speaker's language, and thus to begin to give a truth theory for that language. These axioms will enable him to provide an interpretation of the speaker's non-occasion sentences. However, the range of truth-theories that yield correct truth conditions for the speaker's occasion sentences will at this point be indefinitely large, and the radical interpreter must accept a third principle if he is to narrow down the range of acceptable interpretations. This principle says that the speaker's beliefs are largely coherent. Thus, interpretations that involve the attribution of coherent beliefs to the speaker are to be preferred.

---

<sup>8</sup> For the sake of clarity, I will use male pronouns to refer to the interpreter and female pronouns to refer to the speaker being interpreted.

<sup>9</sup> It may be that some more sophisticated variant on this principle is needed; see Ludwig and Lepore (2005, pp. 182–192) for discussion. However, the version of the principle that I have given here will be sufficient for our discussion.

At this point, pressure to interpret the speaker so as to make her beliefs coherent may lead the interpreter to overturn some of his initial conclusions about the truth conditions of occasion sentences in the speaker's language. In some instances it will greatly increase the overall coherence of the speaker's beliefs if she is interpreted as expressing a false belief with an occasion sentence. Thus, an interpreter can attribute false beliefs about her immediate environment to the speaker. However, whilst concern for coherence might move the interpreter to conclude that any particular occasion sentence expresses a false belief, the interpreter cannot conclude that most or all of the beliefs that the speaker expresses with occasion sentences are false, on pain of undermining the grounds for his whole interpretation.

The three principles just described are usually represented as assumptions that a radical interpreter must make if he is to be able to interpret a speaker at all. It is correct that Davidson describes the radical interpreter as relying on these three principles. However, I do not think that it is correct to see these principles as *assumptions* on the part of the interpreter. In Sect. 2 I will argue that all three principles are implied by a more fundamental principle that guides interpretation, and that it is this fundamental principle that deserves to be labelled the principle of charity.

The Davidsonian response to scepticism may now be put as follows. Davidsonian interpretivism entails that, necessarily, the beliefs of any believer can be discovered by a radical interpreter. The method of radical interpretation requires that the radical interpreter must see the believer as having a lot of true beliefs about her environment. So, any believer must have a lot of true environmental beliefs, and sceptical scenarios are impossible.

So put, it may seem obvious that Davidson's argument is invalid. Stroud (2000, pp. 186–187) writes of the argument I have just sketched that

what we have seen so far to be 'guaranteed' or 'methodologically enforced' by the nature or conditions of interpretation is that if a successful interpreter believes that speakers believe that p, that speakers believe that q, that speakers believe that r, and so on, then in general or in large part that interpreter will also believe that it is true that p, true that q, true that r, and so on. He and the speakers he interprets will be largely in agreement. But that what he and those speakers all agree about is in fact in general true is a further step beyond the fulfilment of that admitted condition.<sup>10</sup>

This is the objection from the possibility of shared error. According to this objection Davidson's interpretivism ensures that the speaker and the interpreter must agree, but it does not thereby ensure that the speaker must have mostly true beliefs, because the radical interpreter himself might have mostly false beliefs.

We can dramatize this idea by imagining that both the interpreter and the speaker are brains-in-vats, wired together so that they experience the same hallucination. One brain might interpret the other as being largely in agreement with himself: he will see the second brain as believing that she has hands, that she is on Earth, that it is sunny, and so on, whilst he himself will believe that she has hands, that she is on Earth, that

<sup>10</sup> See also Stroud (2011, p. 267) for essentially the same point.

it is sunny, and so on. However, what they agree on will be mostly false. Versions of this objection can be found in Williams (1988, p. 189), Ludwig (1992, p. 321), and Brueckner (2010, pp. 176–179).

The objection depends upon our taking it that the second principle of interpretation requires only that the interpreter attribute to the speaker beliefs that agree with her own, rather than beliefs that are actually true. Davidson's claim that the interpreter must 'read some of *his own standards* of truth into the pattern of sentences held true by the speaker' certainly suggests that he takes agreement to be sufficient.<sup>11</sup> If, however, despite this quotation, the second principle requires that the interpreter attribute beliefs to the speaker that are on the whole *actually* true—that is, if the principle requires that the interpreter must on the whole identify the truth conditions of the speaker's occasion sentences with the conditions that *actually* typically cause them—then it is not possible to follow the second principle and ascribe to the speaker mostly false beliefs about her immediate environment. Thus, the objection from the possibility of shared error cannot arise. Let us call the version of interpretivism which interprets the second principle as requiring only agreement between interpreter and speaker agreement interpretivism. We may call the version of interpretivism which interprets the second principle as requiring the attribution of true beliefs to the speaker truth interpretivism.

If the Davidsonian response to scepticism is based upon agreement interpretivism it will require some supplementary argument to meet the objection from the possibility of shared error. It is usually thought that Davidson (1986, p. 317) attempts to provide this supplementary argument in the following passage:

Imagine for a moment an interpreter who is omniscient about the world, and about what does and would cause a speaker to assent to any sentence in his (potentially unlimited) repertoire. The omniscient interpreter, using the same method as the fallible interpreter, finds the fallible speaker largely consistent and correct. By his own standards, of course, but since these are objectively correct, the fallible speaker is seen to be largely correct and consistent by objective standards.<sup>12</sup>

Let us call the argument that this passage is taken to express the omniscient interpreter argument. Roughly, this argument says that an omniscient interpreter can interpret any possible speaker, but this omniscient interpreter will employ the method of radical interpretation, and so will find any speaker he interprets to have empirical beliefs that, on the whole, agree with his own.<sup>13</sup> Since the omniscient interpreter's beliefs are, by stipulation, all true, this entails that most of the speaker's empirical beliefs are true. So, it is not possible for a speaker to have mostly false empirical beliefs.

This perceived move on Davidson's part to ensure that interpretability entails largely true beliefs is widely thought to be a failure. Perhaps the most commonly made objec-

<sup>11</sup> My emphasis.

<sup>12</sup> Likewise, it is usually thought that the same argument is being made by Davidson in 'The Method of Truth in Metaphysics', (1984, p. 201).

<sup>13</sup> The premise must be that any *possible* speaker is interpretable, rather than that any actual speaker is interpretable, to preserve the modal status of the conclusion of the argument.



tion is that the omniscient interpreter argument begs the question against the sceptic.<sup>14</sup> This point can be brought out by considering the idea of an “omni-ignorant” interpreter with entirely false beliefs. Combining the idea that any possible speaker can in principle be interpreted by the omni-ignorant interpreter with the idea that interpretation requires agreement, we can conclude that any being with beliefs must have mostly false beliefs.

If Davidson’s response to scepticism is to succeed we must have some reason for preferring the omniscient interpreter argument over the omni-ignorant interpreter argument. However, as Ludwig (1992, p. 327) says,

it is clear that to have grounds for rejecting as false the assumption that there’s nothing absurd about the idea of an [omni-ignorant] interpreter we would already have to have an argument to show that we can’t be massively mistaken about the world. In assuming, then, that the right premise to start with is the idea that there is nothing absurd in the idea of an omniscient interpreter, we are assuming what we want to prove.

There are other objections to the omniscient interpreter argument, some of which may be correct.<sup>15</sup> However, I am happy to concede that this objection alone is enough to show that the omniscient interpreter argument does not work. If the Davidsonian response to scepticism depends upon the omniscient interpreter argument, then it is a failure.

## 2 Interpretivism and rational action

I will now argue that Davidson subscribes to truth interpretivism, rather than to agreement interpretivism, and so his anti-sceptical argument is immune to the objection from the possibility of shared error. This can be seen by paying attention to what Davidson says about how a theory of interpretation is to be tested in papers that are not directly concerned with scepticism.<sup>16</sup> Indeed, it may be that the popularity of the objection from the possibility of shared error is due to the fact that discussions of Davidson’s anti-sceptical argument do not tend to look beyond his papers on scepticism for their understanding of the interpretivist view on which the argument depends. This is unfortunate, for in his work on scepticism Davidson only sketches his interpretivism, leaving fuller development of the view for elsewhere.

<sup>14</sup> This objection appears in Ludwig (1992, p. 327), Ludwig and Lepore (2005, pp. 328–329), Williams (1988, p. 190) and Dalimiya (1990, pp. 90–94).

<sup>15</sup> See for example Stroud (2000, pp. 187–190), Vermazen (1983), Foley and Fumerton (1985), Ludwig and Lepore (2005, p. 328) and Williams (1988, pp. 190–191).

<sup>16</sup> It is worth noting that consideration of the test of an interpretive theory is not the only way of getting to the exegetical point that, for Davidson, a radical interpreter must have true beliefs about the speaker’s environment if she is to follow the method of interpretation correctly. The same point could, for example, have been made by consideration of Davidson’s (2005, pp. 47–62), (1986, pp. 313–316), (1991b, pp. 158–160) repudiation of the Quineian idea of stimulus meaning.

In ‘Thought and Talk’ Davidson (1984, p. 161) writes that ‘the interlocking of the theory of action with interpretation will emerge ... if we ask how a method of interpretation is tested. In the end, the answer must be that it helps bring order into our understanding of behaviour.’ Ultimately, the interpretation given by a radical interpreter must fit with a general theory that explains some of the speaker’s behaviour as action. That is, some of the speaker’s behaviour is explained by talking about the belief-desire pairs that cause it.<sup>17</sup> That theory must be one on which the actions of the speaker are largely rational: ‘we weaken the intelligibility of attributions of thoughts of any kind to the extent that we fail to uncover a consistent pattern of beliefs and, finally, of actions, for it is only against a background of such a pattern that we can identify thoughts.’<sup>18</sup> The fundamental principle governing interpretation, then, is that the interpretation must fit with an explanation of the speaker’s behaviour, including her linguistic behaviour, as the actions of a rational animal.

Crucially, for Davidson an explanation of behaviour as rational action is not just an explanation of the movement of a part of someone’s body. Rather, it is an explanation of why someone moved their body in such-and-such a way in such-and-such an environment. Take Davidson’s example of flipping a switch as an action.<sup>19</sup> What is wanted here is not an explanation of why someone moved his hand in a certain way; we want an explanation of why he moved his hand into contact with something in his environment (the switch) causing that thing to do something (flip). An example of such an explanation might be that he wanted the light to go on, and that he believed that flipping the switch would result in the light going on.

What goes for explanation of action in general goes also for explanations of actions that are utterances of occasion sentences. The explanandum does not just consist of a speaker’s production of a certain sound at certain times. Rather, the interpreter wishes to explain the fact that certain environmental conditions typically cause the speaker to assent to certain sentences in her language. Thus, the interpreter who follows the method of radical interpretation correctly cannot be mistaken about what circumstances typically cause the speaker to assent to certain occasion sentences. In the process of explaining such facts, the interpreter will apply the second principle of interpretation, identifying the truth conditions of the speaker’s occasion sentences with the conditions that typically cause the speaker to assent to those sentences. Since the interpreter who follows the method of radical interpretation correctly is correct about what these conditions are, he will for the most part assign truth conditions to the speaker’s occasion sentences, and thus to the environmental beliefs that they express, that actually obtain when the speaker assents to those sentences.

The point here can be illustrated by returning to our previous example. The explanandum will not be the mere fact of the speaker’s utterances of ‘*lá se vai um gato*’ at various times. Rather, the explanandum is the fact that *the presence of a cat typically causes the speaker to utter ‘lá se vai um gato’*. Since this is the fact that needs to be explained, the interpreter who is required to explain it must believe it; that is,

<sup>17</sup> Davidson, ‘Actions, Reasons, and Causes’, (1980, pp. 3–20).

<sup>18</sup> Davidson, ‘Thought and Talk’, (1984, p. 159).

<sup>19</sup> Davidson, ‘Actions, Reasons, and Causes’, (1980, pp. 4–5).

the interpreter must believe that the presence of a cat typically causes the speaker to utter 'lá se vai um gato'. By the first principle, the interpreter is to see this sentence as expressing one of the speaker's beliefs. By the second principle, the interpreter will assign to this sentence, and thus to the belief it expresses, the truth condition that there is a cat around. Thus, the sentence and the belief it expresses will have truth conditions that obtain on most of the occasions when the speaker will assent to that sentence, and so also on most of the occasions when she holds the belief it expresses.

Because the interpreter must have correct beliefs about the conditions that typically cause a speaker to assent to occasion sentences it is not possible for an interpreter with mostly false beliefs about the speaker's environment to follow the method of radical interpretation correctly. This is because an interpreter with false beliefs about the speaker's environment will be wrong about what needs to be explained. An interpretation offered by a massively mistaken interpreter may be an explanation of the speaker's actions in what the interpreter *believes* to be the speaker's environment, but it will not explain the speaker's behaviour in her *actual* environment, and so it will not count as an instance of successful interpretation. Similarly, someone with false beliefs about yesterday's events will not be able to follow the instruction: explain yesterday's events. Such a person might provide a good explanation of what they take yesterday's events to be, but this will not be an explanation of yesterday's events. Since it is a requirement of doing radical interpretation correctly that the interpreter have true beliefs about the speaker's environment, the objection from the possibility of shared error does not arise.<sup>20</sup>

I have now given my argument that what Davidson says about the ultimate test of an interpretive theory shows that he endorses truth interpretivism, rather than agreement interpretivism, and so his anti-sceptical argument is not subject to the objection from the possibility of shared error. However, there is more to be said in favour of attributing truth interpretivism to Davidson. Once we see the ultimate aim of interpretation as explanation of the speaker's behaviour in her environment as rational action, we need no longer see Davidson as asserting that the interpreter must simply assume whatever principles are necessary to make interpretation possible. Rather, the principles stem from the overarching aim of explaining the speaker's behaviour in her environment as rational. The speaker's acts of assent to an occasion sentence in certain circumstances are explained as rational because they express the speaker's belief that those circumstances obtain, and this belief fits with the speaker's other beliefs. As Davidson puts it, 'each interpretation [of a sentence] and attribution of attitudes is a move within a holistic theory, a theory necessarily governed by concern for consistency and general coherence with the truth.'<sup>21</sup>

It may be that there is a "thin" sense of rationality that does not require 'general coherence with the truth', but only coherence of beliefs. Even a brain-in-a-vat, conceived of as having almost entirely false beliefs, may be rational in this sense. Be that as it may, the point here is that the interpreter has no choice but to think of the

---

<sup>20</sup> Manning (1995, p. 344) can be read as suggesting that it is a condition of doing radical interpretation correctly that the interpreter have true beliefs about the speaker's environment. However, he does not explain *why* this is a condition of doing correct interpretation.

<sup>21</sup> Davidson 'Belief and the Basis of Meaning' (1984, p. 154).

speaker as being rational in a “thick” sense, which does require ‘general coherence with the truth’. The assumption that the speaker is rational in the thick sense allows the interpreter to identify the truth conditions of the speaker’s occasion sentences with the conditions that typically cause them, thus allowing the process of interpretation to get started. Without it, there is no way for the interpreter to begin.

The idea that the interpreter need only see the speaker as rational in the thin sense merits further attention, since if the interpreter need only see the speaker as rational in the thin sense Davidson’s anti-sceptical argument is completely undercut. If the method of radical interpretation only required the interpreter to see the speaker as rational in the thin sense it would not require the interpreter to ascribe beliefs to the speaker that agree with his own, as agreement interpretivism requires, or that are actually true, as truth interpretivism requires. It would require only that the interpreter ascribe coherent beliefs to the speaker. There would be no requirement that these beliefs would be true, or even that they would be thought to be true by the interpreter.

The problem with the idea that the interpreter is only required to interpret in a way that makes the speaker rational in a thin sense is that there are many different interpretations of the speaker’s language which would involve the attribution of coherent but mutually exclusive belief systems. Without the assumption that most of the speaker’s occasion sentences are true the interpreter seems to have no grounds for preferring one of these interpretations over any of the others. This might be thought to be too quick, however. Could not the interpreter narrow down the range of interpretations by preferring those interpretations that fit with the way things *seem* to the speaker? For example, the interpreter provisionally concludes that the speaker’s utterances of ‘lá se vai um gato’ are true if and only if there is a cat around because she typically assents to this sentence when it *seems to her* that there is a cat around.

The problem with this suggestion is that it fails to take seriously the idea that the evidence upon which an interpretation is based must be available from the third person perspective of the radical interpreter. Its seeming to the speaker that such-and-such is the case is a matter of the speaker’s having certain subjective experiences. But facts about the speaker’s subjective experiences will not be directly available to the interpreter. The interpreter cannot, for example, observe the sensory experiences that constitute its seeming to the speaker that there is a cat in front of her, as he can observe that there *is* a cat in front of her. Thus, the interpreter cannot observe that the speaker’s having certain subjective experiences correlates with her utterances of ‘lá se vai um gato’. Of course, the interpreter may eventually be able to make successful inferences about the speaker’s subjective experiences, but it is hard to see how he could come to know when it seems to the speaker that a cat is around without first arriving at an interpretation of the speaker’s language.

I will end this section by considering some passages from Davison that might be thought to cause trouble for the idea that he is a truth interpretivist. First, as we have seen, Davidson (1986, p. 317) responds to the worry that ‘the speaker and interpreter understand one another on the basis of shared but erroneous beliefs’ by asking us to imagine an interpreter who is ‘omniscient about the world, and about what does and would cause a speaker to assent to any sentence in his (potentially unlimited) reper-

toire'.<sup>22</sup> What are we to make of this reply if it is not an expression of the omniscient interpreter argument, intended to plug the hole in an anti-sceptical argument founded upon agreement interpretivism?

I suggest that the real point of the passage is not that an omniscient interpreter is possible. Rather, the point is what follows from that omniscience: correct beliefs about 'what does and would cause a speaker to assent to any sentence in his (potentially unlimited) repertoire'. Davidson is simply pointing out that the method of interpretation requires an interpreter to have correct beliefs about these things to follow the method of radical interpretation correctly, and so interpretation on the basis of shared but false beliefs 'cannot be the rule'. He is not adding the (question begging) claim that it is in principle possible for any speaker to be interpreted by an omniscient interpreter to his claims about the method of radical interpretation in an effort to make his response to scepticism sound.<sup>23</sup>

There is also the fact that Davidson frequently writes in a way that seems to suggest that all that an interpreter needs to do to interpret correctly is to take a speaker to have beliefs that agree with her own. For example, I took Davidson's assertion that the interpreter must 'read some of his own standards of truth into the pattern of sentences held true by the speaker' as a gloss on the second principle of interpretation.<sup>24</sup> Similarly, in 'Mental Events' Davidson (1980, p. 222) writes that when we interpret 'we will try for a theory that finds [the speaker to be] consistent, a believer of truths, and a lover of the good (all by our own lights, it goes without saying)'. How should we explain these quotations if the method of radical interpretation requires an interpreter to attribute true beliefs to a speaker?

Consider claims of the form: necessarily (in aiming to  $\phi$ , one will  $\psi$ ). It does not follow from a claim of this form that one's aim to  $\phi$  is satisfied by  $\psi$ -ing. Consider again the example of aiming to explain yesterday's events. In doing so I will necessarily explain what I take yesterday's events to be. However, it does not follow that my aim of explaining yesterday's events is fulfilled by explaining what I take yesterday's events to be. My aim is only fulfilled if I explain what actually happened yesterday.

This distinction between what one is aiming to do and what is necessary to achieve one's aim explains the way in which Davidson writes. In attempting to explain the facts about a speaker's behaviour and utterances in her environment an interpreter can do no better than go by what she believes the facts about the speaker's behaviour and utterances in her environment to be (this is what 'goes without saying'). Thus, in aiming to interpret the speaker's language so that most of her occasion sentences express true beliefs the interpreter will, necessarily, read 'some of his own standards of truth into the pattern of sentences held true by the speaker'. The method of radical interpretation does require agreement. That is why Davidson writes as he does, and why people have made the mistake of seeing him as an agreement interpretivist. However, radical interpretation requires more than just agreement. Only insofar as he manages

<sup>22</sup> This is part of the passage quoted at the end of Sect. 1.

<sup>23</sup> Ludwig and Lepore (2005) also suggest that Davidson does not endorse the omniscient interpreter argument, and report Davidson as saying in conversation that he wished he had never mentioned the omniscient interpreter.

<sup>24</sup> Davidson 'A Coherence Theory of Truth and Knowledge', (1986, p. 316).

to interpret the speaker's language so that her occasion sentences express true beliefs, and thus, ultimately, explains the speaker's behaviour in her environment as rational action, does the radical interpreter satisfy her aim of following the method of radical interpretation.

### 3 Truth interpretivism versus agreement interpretivism

So far, I have been concerned to establish the exegetical conclusion that Davidson endorses truth interpretivism rather than agreement interpretivism. If Davidson's truth interpretivist view is true, then sceptical scenarios are impossible. The next question for anyone evaluating Davidson's anti-sceptical argument will of course be whether the antecedent of this conclusion is true. This question is well beyond the scope of this paper. However, we are in a position to see that truth interpretivism is preferable to one obvious competitor, namely agreement interpretivism.

Agreement interpretivism allows that both the omniscient interpreter and the omniscient interpreter may follow the method of radical interpretation correctly. Each interpreter can interpret the speaker in a way that attributes beliefs that largely agree with his own. Yet, each interpreter will end up with very different conclusions about what the speaker means and believes. (Indeed, there is nothing in agreement interpretivism to prevent many different omniscient interpreters starting out with different sets of false beliefs about the speaker's environment from following the method of radical interpretation correctly, and ending up with different conclusions about what the speaker means and believes.)

The conclusions reached by the omniscient interpreter and the omniscient interpreter will be logically incompatible. Suppose that the omniscient interpreter concludes that the speaker believes that she has hands. The omniscient interpreter does not believe that the speaker has hands (since it is true that the speaker has hands). So the omniscient interpreter concludes that the speaker does not believe that she has hands. However, the agreement interpretivist must admit that both of these interpretations are correct. Each interpreter has followed the method of radical interpretation (as the agreement interpretivist conceives it to be) correctly. And as an interpretivist, the agreement interpretivist thinks that any conclusion reached about what a speaker believes reached by following the method of radical interpretation correctly must be correct. So it seems that the agreement interpretivist faces a contradiction: the speaker believes that she has hands, and it is not the case that the speaker believes that she has hands.

Truth interpretivism, on the other hand, does not fall into contradiction in this way. The problem arises for agreement interpretivism because it allows for the possibility that both an omniscient and an omniscient interpreter might correctly apply the method of interpretation to the same speaker. The truth interpretivist, on the other hand, can disqualify the interpretation offered by the omniscient interpreter on the grounds that it does not explain some of the speaker's behaviour in her environment as rational action, thus avoiding contradictions of the sort just described. This is one good reason to prefer Davidson's truth interpretivism over the agreement interpretivism that is erroneously ascribed to him.

## 4 Conclusion

The main purpose of this paper has been to defend Davidson's anti-sceptical argument from the objection from the possibility of shared error. This objection arises when we take Davidson's argument to rest on agreement interpretivism, which requires only agreement between interpreter and speaker, thus leaving open the possibility that interpreter and speaker are in a shared state of massive error.

In fact, however, Davidson's anti-sceptical argument rests on truth interpretivism, which requires the interpreter to explain a speaker's acts of assent to occasion sentences in the circumstances that actually typically cause them (as opposed to the circumstances that the interpreter believes typically cause them) as instances of rational action. This requires the interpreter to interpret most of the speaker's sentences as being true in the circumstances that typically cause the speaker to assent to them, and thus to ascribe to the speaker mostly true beliefs about her environment. Thus, it is not possible for an interpreter with mostly false beliefs about the speaker's environment to follow the method of radical interpretation correctly, and the objection from the possibility of shared error does not arise. Davidson's interpretivist view does entail that sceptical scenarios in which thinkers have mostly false beliefs about their environment are impossible.

I hope that by showing that Davidson's anti-sceptical argument is not vulnerable to the objection from the possibility of shared error I have done something to revive interest in the argument among epistemologists. Once it has been accepted that the argument is not simply invalid a number of interesting questions arise. One question will of course be whether the truth interpretivist view on which the argument depends is plausible. I have argued that it is preferable to the agreement interpretivist view erroneously attributed to Davidson by those who make the objection from shared error. In so doing I hope to have done something to recommend truth interpretivism to those who are already attracted to an interpretivist view of meaning and belief. However, it would be preferable to give a more widely appealing argument for truth interpretivism. Davidson himself gives several arguments for his interpretivist view, the most explicit of which appear in his later work involving the idea of triangulation. I hope that this paper has shown that these arguments should be of interest to epistemologists, as well as to philosophers of mind and language.

Another question is, should Davidson's anti-sceptical argument turn out to be sound, what are we to make of the intuitive idea that one might be in a brain-in-a-vat with mostly false beliefs? This strikes most people as being at least a logical possibility, and one might think that there is some pressure for anyone who thinks they have an argument that it is not to provide an explanation of why so many people have made this mistake. I believe that such an explanation should start by pointing out that it is not entirely a mistake. Davidson's argument does not entail that one could not be a brain-in-a-vat. It only entails that one could not be a brain-in-a-vat with mostly false beliefs. The logical possibility that one might be a brain-in-a-vat with mostly true beliefs about one's environment, or with no beliefs at all, remains open. So the answer must focus not on explaining why people are mistaken in assuming that it is logically possible that they are brains in vats, but on explaining why people mistakenly think that if they were brains in vats they would have mostly false beliefs. However, an attempt

to reconcile our intuitions about sceptical scenarios with a Davidsonian response to scepticism must be postponed for another occasion.

**Acknowledgements** My thanks to Bernhard Salow, Crispin Wright, Peter Sullivan and Adrian Haddock for helpful discussion, and to two anonymous Synthese referees for comments that have greatly improved the paper. I am also grateful for a research grant from the São Paulo Research Foundation (Grant ID No. 2016/03277-1) that allowed me to complete this paper.

## References

- Brueckner, A. (2010). *Essays on scepticism*. Oxford: Oxford University Press.
- Byrne, A. (1998). Interpretivism. *European Review of Philosophy*, 3, 199–223.
- Childe, B. (1994). *Causality, interpretation, and the mind*. Oxford: Oxford University Press.
- Dalmiya, V. (1990). Coherence, truth and the omniscient interpreter. *The Philosophical Quarterly*, 40(158), 86–94.
- Davidson, D. (1980). *Essays on actions and events*. Oxford: Oxford University Press.
- Davidson, D. (1984). *Inquiries into truth and interpretation*. Oxford: Oxford University Press.
- Davidson, D. (1986). A coherence theory of truth and knowledge. In E. Lepore (Ed.), *Truth and interpretations* (pp. 307–319). Oxford: Wiley.
- Davidson, D. (1991a). Epistemology externalized. *Dialectica*, 2(3), 191–202.
- Davidson, D. (1991b). Three varieties of knowledge. In A. P. Griffiths (Ed.), *A.J. Ayer memorial essays* (Vol. 30, pp. 153–166). Cambridge: Cambridge Press.
- Davidson, D. (2005). *Truth, language, and history*. Oxford: Oxford University Press.
- Dennett, D. (1971). Intentional systems. *Journal of Philosophy*, 68, 87–106.
- Foley, R., & Fumerton, R. (1985). Davidson's theism? *Mind*, 48, 83–89.
- Lewis, D. (1974). Radical interpretation. *Synthese*, 27(3/4), 331–344.
- Ludwig, K. (1992). Scepticism and interpretation. *Philosophy and Phenomenological Research*, 52(2), 317–339.
- Ludwig, K., & Lepore, E. (2005). *Meaning, truth, language, and reality*. Oxford: Oxford University Press.
- Manning, R. N. (1995). Interpreting Davidson's omniscient interpreter. *Canadian Journal of Philosophy*, 25(3), 335–374.
- Stroud, B. (2000). *Understanding human knowledge*. Oxford: Oxford University Press.
- Stroud, B. (2011). *Philosophers past and present: Selected essays*. Oxford: Oxford University Press.
- Vermazen, B. (1983). The intelligibility of massive error. *The Philosophical Quarterly*, 33(130), 69–74.
- Williams, M. (1988). Scepticism and charity. *Ratio*, 1(2), 34–66.