



Radical Interpretation Interpreted

Donald Davidson

Philosophical Perspectives, Vol. 8, Logic and Language (1994), 121-128.

Stable URL:

<http://links.jstor.org/sici?sici=1520-8583%281994%298%3C121%3ARII%3E2.0.CO%3B2-A>

Philosophical Perspectives is currently published by Blackwell Publishing.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/black.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

RADICAL INTERPRETATION INTERPRETED

Donald Davidson
University of California, Berkeley

I'm sorry that this paper, "Is Radical Interpretation Possible?" (henceforth RIP), does not address the question its title asks. I wish it did; it is a good question. I can't blame Jerry Fodor and Ernie Lepore for suspecting that "it may turn out that perfectly kosher languages (like, for example, English) *aren't* radically interpretable"; I share the suspicion. But unfortunately they never discuss, much less demonstrate the inadequacy of, the fairly elaborate account I have developed over the years of how I think it might be done. Instead they demolish, to their own satisfaction, an argument they think is mine, and then "deny that any other plausible reason" has been given for holding that natural languages are radically interpretable. Since the argument they claim to have demolished is not mine, it is a pity they did not explain why they consider the reasons I do give implausible.¹

RIP is largely devoted to attacking arguments of what its authors call "form T". As far as I know, I have never endorsed any argument of form T; the authors of RIP refer to no text of mine that contains such an argument. Let me try to make clear the relation of my views to arguments of form T.

Premise 1. I do not think I have ever argued for the claim that radical interpretability is a condition of interpretability.² Not only have I never argued that every language is radically interpretable; I have not even argued that every language can be understood by someone other than its employer, since it would be possible to have a private code no one else could break. I do not think, and have not argued, that radical interpretation of natural languages *must* be possible; I have argued only that it *is* possible. The point of the "epistemic position" of the radical interpreter is not that it exhausts the evidence available to an actual interpreter, but that it arguably provides sufficient evidence for interpretation.

Premise 2. What, exactly, are we to put for "F"? The only substitution instance Fodor and Lepore ascribe to me is "the principle of charity". It is true

that I believe that all successful interpretation depends upon the application of the principle of charity, and so, of course, that radical interpretation depends on it. But I do not believe I have presented any argument that makes the need for charity depend on *radical* interpretation; the argument goes the other way around. Fodor and Lepore think Wittgenstein may have argued that there can't be a private language along the lines of argument form T. It's unclear to me how to fit anything Wittgenstein says into an argument of this form, but fortunately I don't have to defend Wittgenstein. I have maintained that anyone with *any* language must have a non-private language (by which I mean a language more than one person understands), but my argument doesn't follow any format Fodor and Lepore discuss.

I have not maintained anything like premise 2 because I think various *stronger* claims, claims of the sort that might be expressed by versions of the conclusion, 4, are evident in themselves. Rather than argue for (some form of) premise 2 I have argued *independently* (without appeal to anything like premise 2) that all interpretation, whether radical or not, must be constrained in certain ways, and therefore that all natural languages must have certain properties. Clearly, nothing here fits argument form T, since I am, among other things, arguing for various conclusions about the nature of (first) languages without appeal to premises like 1 and 2.

Premise 3 (and first conclusion). Fodor and Lepore's third item contains both a premise and a conclusion: the premise is that natural languages are actually interpreted; the conclusion (drawn from this premise and premise 1) is that natural languages are radically interpretable. It should already be clear that I have no use for this step. It is not the fact that natural languages are (in some sense) actually interpreted that has persuaded me that radical interpretation is possible; it is the account I have given of *how* it could be done. I say a little more below about my views on the actual interpretability of natural languages; absent these views I would not find my account of radical interpretation very interesting.

4. Second Conclusion. Some of the theses Fodor and Lepore would include here are theses I have maintained, but not on the grounds they list or discuss. The general format of the strategy (this can't really be called an argument) I have followed with respect to items of the sort that feature in Fodor and Lepore's Argument "Form T" is something like this:

- A. Natural languages have certain properties. (I list some of these below; the arguments don't depend on radical interpretation.) These properties place constraints on answers to the question how coming to understand a language without the aid of a bilingual is possible.
- B. Radical interpretation is a way someone could come to understand a natural language within the constraints mentioned in A.

C. So someone could come to understand a natural language (and we see one way how).

Fodor and Lepore's conclusion (4) is represented here by the first premise, A. B is the claim they question in their title, and which turns up in the conclusion included in their premise 3. (But I don't argue for B from any of the premises they list.) Aside from the parenthetical remark, C is something we knew to begin with, so the parenthetical remark explains (whether or not it justifies) my interest in radical interpretation.

Here are some of the properties of natural languages mentioned in A for which I have argued:

A1. A learnable (or interpretable) language must have a finite primitive vocabulary. The argument is that we are not born knowing the languages we come to speak, so we must learn the meaning (interpretation) of each independent item in the vocabulary. Given that it takes a finite length of time to learn each item, and that men and women are mortal—well, you see how it goes. The point may be trivial, but a number of semantic theories have failed to observe it. In any case, Fodor and Lepore apparently don't dispute this.

A2. Because sentences have propositional contents, sentences have logical properties. A speaker can't be conveying (in the ordinary sense—knowing what his words mean) a particular proposition with his words if he has no idea what properties that proposition has. This does not need to be put in terms of explicit knowledge: a given sentence in someone's language can't express a certain proposition unless it occupies a place among other sentences that reflect the logical relations among those sentences. Thus no one can believe that $p \& q$ and not believe that p ; no one can regularly use a sentence to assert that $p \& q$ when that person is not prepared to assert that p . (This is clearly not an ordinary empirical claim.)

If it is accepted, this thesis (A2) places a restriction on interpretation (*any* interpretation, from any epistemic stance): it makes no sense to interpret some linguistic device as conjunction which the users of that device do not generally use in accord with the truth table for conjunction. Interpretation depends on (cannot avoid) reading some of the norms of the interpreter into the actions and speech of those he interprets. This is one facet of the principle of charity.

I'm not supposing these brief remarks can do more than remind those already familiar with the position that there are *arguments* here; the arguments may be wrong, but they aren't silly, and can't be dismissed without being faced.

A3. I have argued that many basic sentences must be true at those times when they are held true by a speaker. This thesis, like A1 and A2,

is certainly relevant to the question of the possibility of radical interpretation, but since Fodor and Lepore do not discuss my reasons for holding it, there is no point in repeating my (readily available) arguments here.

A4. In my view, it follows from A2 that what one sentence means depends on the meanings of other sentences. I am not an unbuttoned holist in that I do not say the meaning of a sentence depends on the meanings of all sentences.³ Holism, however it is construed, does not follow from the epistemic constraints on radical interpretation: radical interpretation is constrained by holism.

If A2 and A3 are correct, then arriving at an understanding of a speaker's language is legitimately held to be different from arriving at a theory in physics or chemistry. The reason is that in physics and chemistry we don't treat electrons and quarks as having thoughts. Atoms and black holes don't care if we understand them. When people talk or size each other up, there are minds at work at both ends.

Now I turn to a fuller consideration of premise 1 of RIP's argument form T, for I sense that what Fodor and Lepore find most questionable about radical interpretation is the constraints that they think I have, without good reason, put on it and it is here that they have most woefully misunderstood and misread me. Let me list the ways.

1. I have never claimed to know how children learn their first language. (In fact, it is a mystery to me how we can correctly describe the contents of a partly formed mind.)

2. I have never claimed to give an account of how field linguists arrive at their theories.

3. I have never suggested, and in fact have explicitly denied, that the evidence I argue would *suffice* to confirm an interpretive theory exhausts the evidence actually available to interpreters. My argument that radical interpretation is possible therefore does not depend, and was never claimed to depend, on pretending that the evidence and methods I describe represent the "actual epistemic condition of children learning a first language or of linguists and translators in the field". The case I have made for the possibility of radical interpretation does not depend on any assumption about the totality of evidence available to a field linguist.

4. Nothing I have said limits the general intellectual resources of an interpreter. It's true that I don't make use of the powerful knowledge Fodor and Lepore say linguists have at their disposal about linguistic universals, what theoretical devices have proven successful, etc., if only because people without these resources seem able to learn alien languages (and maybe I suspect that linguists, like philosophers, are in some "disarray" in these matters). But this is

not to the point, since I am outlining what I claim *could* succeed, not what *does*.⁴

But even my theoretically bereft interpreter has a lot more going for her than Fodor and Lepore seem to think. I have stressed that a radical interpreter already has a language, and a set of concepts that more or less match those of the interpretee. My interpreter has the concepts of truth, of intention, of belief, of desire, and of assertion (and many, many more). She knows a lot about the world and about how people behave in various circumstances. I have argued in some detail that interpretation can succeed only when interpreter and interpretee are much alike in important respects (alike in what they can perceive, alike with respect to built-in and learnable patterns of discrimination, alike even in size and degree of mobility).

What *doesn't* my radical interpreter know? Well, she's not a mind-reader, and she hasn't learned what someone thinks or means by opening up his brain. The important limitation is that she doesn't know in detail the contents of any of the propositional attitudes of the person to be interpreted: she doesn't know what he intends, believes, wants or means by what he says. (The *reason* for this limitation is philosophical, not scientific, and has to do with my actual interest in radical interpretation. I shall say more about this presently.)

How about the "nothing is hidden" thesis? So far as the defense of the possibility of radical interpretation is concerned, I have certainly intended to allow the radical interpreter only evidence that is not hidden. The authors of RIP say I "stipulate" this because they think I believe this is all the evidence anyone actually has; but as I have said, they are wrong about my views on this point.⁵ Close at hand, however, is a point on which the authors of RIP and I may well disagree, and this disagreement may partly underlie and explain the extent to which they have misunderstood me. I hold that the speakers of natural languages can be, and often are, correctly understood on the basis of non-linguistic facts not merely available, but readily available, to the likes of you and me. This view does not entail that we may not have a universal grammar wired in, or even that we don't have the grammars of five or fifty thousand languages wired in. But it does embrace the conviction that the references of unstructured singular terms and predicates are not wired in; if we understand a speaker, we know how her words are connected to the world, and this is something we cannot be born with. These connections are established both for the speaker and the interpreter in the context of social interaction. Meaning is not something each of us harbors, and which others may or may not discover, but which could exist forever uncommunicated. It is not important to me to call what those who learn a language must somehow observe "evidence"; what matters to the defense of the possibility of radical interpretation is that a mature observer with a language of her own could treat it as evidence.⁶

So I think Fodor and Lepore are mistaken when they say "*Whatever* one's epistemic situation (short of omniscience), it always *could* turn out that the

available data are misleading.” Of course we can be wrong in our interpretation of any particular person or utterance; but if we never understood anyone, the concepts of language, understanding and thought would have no application to us. Such a person could not be “misled” by evidence. Fodor and Lepore say they “propose not to beg the case against Intentional Realism”. In what immediately follows, however, it is clear that what they mean is that they take such realism for granted. I have no objection if “Intentional Realism” entails no more than that there are correct and incorrect interpretations of what a speaker means by his words; it is hard to think what point there could be to radical interpretation if it did not claim to make this distinction. One can, in my opinion, be a realist in this sense and consistently hold that the primary and only ultimate source of meaning lies in successful interpersonal communication. It is only if we have a Cartesian, individualistic conception of meaning and the intentional that we assume a conflict between realism and holism. Realism about correct interpretation does not, for me, entail that what someone means by his words is independent of what is understood by others, nor does it imply that what expressions in a natural language mean is independent of how its speakers understand one another. When Fodor and Lepore say it always could turn out that the available data are misleading, they seem to be assuming that realism requires such independence. This assumption, for which they do not argue, may explain why they cannot grasp the idea that there may be a fundamental difference between how we come to know what others mean and think and how we come to know, if we do, how the blood circulates or planets are formed.

The authors of RIP think my chief and perhaps only investment in radical interpretation is to prove (“transcendentally”, whatever that may mean here) that natural languages must have certain properties. There they are wrong: I do not see radical interpretability as a Procrustean bed to which language must be made to fit. Why then do I care if radical interpretation is possible? And are they right that if it fails to prove what they and I agree it doesn't prove, then it “has no important role to play in theorizing about language and mind”?

Like many others, I wanted answers to such questions as “What is meaning?”, and became frustrated by the fatuity of the attempts at answers I found in Ogden and Richards, Charles Morris, Skinner and others. So I substituted another question which I thought might be less intractable: What would it suffice an interpreter to know in order to understand the speaker of an alien language, and how could he come to know it? (Where I look mainly at the interpreter, others have asked the analogous questions about a speaker, but for various reasons the difference isn't very important at this degree of abstraction.)

I suggested that a theory of truth, constructed more or less along the lines of one of Tarski's truth definitions, would go a long way toward answering the first question. It was important to me that such theories were real theories that could be axiomatized, their primitive concepts and defined terms unambiguously identified, and theorems proved in and about them. I have always been aware that

it is a big question whether, or to what extent, such theories can be made adequate to natural languages; what is clear is that they are adequate to powerful parts of natural languages, parts with great expressive power. Of course I did not say speakers or interpreters actually formulate such theories. It does seem to me, though, that if we can describe how they *could* we will gain an important insight into the nature of the intentional (including, of course, meaning), in particular into how the intentional supervenes on the observable and non-intentional.

How could someone come to know that such a theory correctly describes a speaker or group of speakers? For reasons at which I hinted above, and have explored at length elsewhere, I believe that what people mean by what they say derives from occasions of successful communication, so I restrict the evidence to what would be plainly available to an observer unaided by instruments. The other important restriction is to assume no prior detailed knowledge of any of the propositional attitudes: beliefs, desires and intentions. As a result, I consider a theory of meaning to be an undetachable part of a more general theory of human behavior.

My strategy can be compared to, and was partly inspired by, Frank Ramsey's strategy in trying to make sense of the concept of subjective probability or degree of belief. Noting that subjective probabilities are not directly observable by others, and cannot generally be evaluated by those who have them, Ramsey contrived a theory in which the important undefined primitive was preference, and showed how to extract subjective probabilities and cardinal utilities from the resulting structure. Unquantified preference, though not defined, is plausibly taken to be revealed by choice behavior; choice behavior, being relatively observable, thus provides the evidence on the basis of which the theory can be given empirical content. The more subtle psychological attitudes of degree of belief and relative strengths of preference (cardinal utilities) are theoretical concepts, given meaning by their roles in explaining and predicting choices. Ramsey outlined a sequence of choices which, given the theory, would yield the desired information about degrees of belief and cardinal utilities. This sequence can be turned into a formal proof of the adequacy of the theory. By analogy, I would take certain observable aspects of verbal behavior as evidence for a theory of truth and, by combining a theory of truth with a version of decision theory, a theory of belief, meaning and desire. The story I tell about how radical interpretation is possible should be viewed as an informal proof that the evidence assumed available plus the constraints on the structure of the pattern of a person's beliefs, values and language, suffice to yield an interpretive theory for understanding that person.

It seems clear to me that the very general doubts Fodor and Lepore have raised do not touch the question whether radical interpretation (as I understand it) is possible. To show this, or even that it is unlikely to work, they would have either to produce *a priori* reasons why you can't get there from here, or come to

grips with the arguments which aim to show in some detail how it could be done.

Notes

1. I would like to enter a mild protest against the slack way in which "the friends of radical interpretation" are lumped together. Who are they? Quine is named, but it's neither his phrase nor his idea. David Lewis, Daniel Dennett, and Richard Grandy are mentioned, but the only one of these who has used the term "radical interpretation" is David Lewis, and that was in replying to a paper of mine. Of course I'm not going to defend all these worthies; they are quite capable of defending themselves. In any case, I don't think Fodor and Lepore do any of them any more damage than they do me. But why list a miscellaneous group of people and accuse them of being in "disarray" because they don't agree on everything? It's like accusing someone of not being a good member of a club he never joined.
2. Reading over the papers I have written on interpretation over the last 27 years, it is clear to me that neither my terminology nor my views have held absolutely steady. I have sometimes used the term "radical interpretation" to refer to any interpretation from scratch, that is, without the aid of bilingual speakers or dictionaries, and sometimes to refer to the special enterprise of interpreting on the basis of a limited and specified data base. Fodor and Lepore seem to restrict the use of "radical interpretation" to the second case; in this reply I cleave to this convention. But as far as I can find, I have never argued, specified, or assumed, as Fodor and Lepore insist, that the data on which the special enterprise is based exhaust the data available to actual interpreters.
3. In "Truth and Meaning", first published 25 years ago, I wrote, "Frege said that only in the context of a sentence does a word have a meaning; in the same vein he might have added that only in the context of the language does a sentence (and therefore a word) have a meaning." (Reprinted in *Inquiries into Truth and Interpretation*, Oxford University Press, 1984, p. 22.) I have subsequently been more restrained on this topic.
4. For a recent statement of my view of the difference between radical interpretation and the ways we actually come to understand one another, see "The Structure and Content of Truth", *The Journal of Philosophy*, 87 (1990), p. 324 f. and note 67.
5. Fodor and Lepore quote a long passage from page 314 of "The Structure and Content of Truth", which endorses the thesis that whatever it is that we have to learn in order to understand others must be based on "readily observable behavior". They think I hold this thesis to be true because of the actual success of radical interpretation. I hope less prejudiced readers of this passage and its context will take it rather as a defense of the strategy of radical interpretation, which depends on such "readily observable behavior".
6. Perhaps I should once more point out that it is because I hold that natural languages are "epistemically accessible" that I require that radical interpretation be based on readily available evidence. So it doesn't matter to the possibility of radical interpretation whether or not I am right about this; if I am wrong, I have merely made life unnecessarily difficult for myself.