The Dead End of Radical Interpretation

Roberto Horácio de Sá Pereira
UFRJ/CNPq

The Quinean Heritage

At a first pass, it might appear misleading to construe Davidson as attempting to provide a theory of meaning on a purely extensional basis. For one thing, Davidson's work is informed by Quine's semantic nihilism couched by his indeterminacy of meaning thesis. For another, Davidson's initial project was to pursue a compositional theory of meaning constrained by a theory of truth à la Tarski and a theory of communication and understanding that avoids reification of meaning, as Church's semantics exemplify, at least according to Davidson.

On closer inspection though, Quine's philosophy of language has an immeasurable impact on Davidson. Quine's philosophy sets the stage for Davison's theory of meaning. The core of Davidson's philosophy of language is Quine's idea of a radical translator, who interprets another's language on the basis of evidence that does not presuppose any detailed knowledge of his thoughts or any knowledge of the meanings of his words. The choice of this as the stance from which to investigate language and thought and their relation to the world is motivated by the flawed assumption that interpretation constitutes thoughts and beliefs. In his own words:

> the third-person approach to language is not a mere philosophical exercise. The point of the study of radical interpretation is to grasp how it is possible for one person to come to understand the speech and thoughts of another, for this ability is basic to our sense of a world independent of ourselves, *and hence to the possibility of thought itself*. The third-person approach is yours and mine. (Davidson 2004 (2001): 143, emphasis added)

This is the main assumption that I want to challenge is this paper. The radical interpreter must construct his picture of another's thoughts and words and his connection with reality all at once out of the public resources available to any objective knower. Out of the investigation of the constraints on radical interpretation and the assumptions needed for success emerges a unified picture of human beings whose mind is constituted by the trade-offs between language and the world. Yet, even by accepting Quine's framework, Davidson rejects Quine's semantic nihilism by rejecting Quine's thesis of indeterminacy: the same foreign sentence can be translated equally well by two (or more) different home language sentences. In the place of Quine's nihilism Davidson proposes a semantic reductionism:

> When all the evidence is in, there will remain, as Quine has emphasized, the trade-offs between the beliefs we attribute to a speaker and the interpretations we give his words. But the resulting indeterminacy cannot be so great but that any theory that passes the tests will serve to yield interpretations' (1984: 139).

> [I]ndeterminacy [of meaning or translation] is important only for calling attention to how the interpretation of speech must go hand in hand with the interpretation of action generally, and so with the attribution of desires and beliefs' (1984: 154).

This is the background to Davidson's philosophical agenda. On the one hand, he acknowledges against Quine the importance of a theory of meaning and rejects Quine's semantic nihilism. On the other, following Quine, he rejects the traditional assumption that intensional notions are primitive, that is, unaccountable in terms of more primitive notions. Moreover, he clearly endorses the highly questionable Quinean assumption that to avoid circularity a theory of meaning must explain communication without relying on intentional or semantic notions. The alternative is to propose a form of semantic reductionism that aims to account for semantic or intentional notions on a purely extensional basis: Tarki's theory of truth. Has such a program any chances of success? The answer is no.

My aim in this paper is not to show that Davidsons fails in addressing the so-called extensionality problem. That is no novelty. Even Davidsonians recognize that "Davidson's initial hope that an extensionally adequate truth theory would *ipso facto* be interpretive fails" (Lepore & Ludwig 2005: 102). Yet, they consider that the extensionality problem is "not fatal to the initial

project of pursuing a compositional meaning theory by way of an appropriately constrained truth theory" (Lepore & Ludwig 2005: 102). In contrast, my aim is to show that Davidson's failure brings us back to Quine's semantic nihilism and also renders his account of meaning something quite trivial. My main goal is to argue that Davidson & Quine's framework, the radical translation/interpretation, is at the root of such failure. There is no priority of language over thought: *radical interpretation is not the basis of sense and hence the possibility of thought.* This is a remnant of the so-called linguistic turn that Davidson inherited from the analytical tradition. We must come back to the commonsense notion that to learn a language and to translate an alien tongue into our own is merely to learn a means of communicating and expressing thoughts, beliefs, or desires that we have already had.

How shall I proceed here? After this brief introduction, the next section is devoted to presenting Davidson's project and its extensionality problem. From the third section onwards, I begin to address the solutions proposed by Davidson for the extensionality problem. The third section, in particular, analyzes Davidson's meaning of holism as a solution to the extensionality problem. I argue that this and the compositionality of natural languages is neither a *necessary nor a sufficient condition* for the extensionality problem. In the fourth section I examine Davidson's principle of interpretative charity as a solution to the extensionality problem. In this section I argue that apart from not solving the problem, Davidson's main formulation of the problem begs questions. In the fifth section I argue that the putative lawlike character of theorems also fail in providing an answer to the extensionality problem since laws are indifferent to co-extensive predicates with different meanings. Here I argue briefly that Davidson's idea of a reductionist theory of meaning can only be carried over if we abandon the methodology of radical interpretation.

The Extensionality problem

Like Quine, Davidson believes that we cannot take any intensional notion as a primitive or given concept. Nonetheless, in opposition to Quine he still believes we can and must explain the concept of meaning without appealing to any intensional notion.

The starting point of a theory of meaning on a purely extensional basis comes to be Tarski's celebrated theory of truth:

A theory of truth will be materially adequate, that is, will correctly determine the extension of the truth-predicate, provided it entails, for each sentence s of the object language, a theorem of the form: "s is true iff p" where 's' is replaced by a description of s and 'p' is replaced by a sentence that is true if and only if s is. (1984: 150)

Davidson's egg of Columbus is this: while Tarski takes the notion of *translation* as primitive and on this basis attempts to explain the notion of *truth*, Davidson reverses the relation between *explicans* and *explicandum*: he takes the notion of truth as primitive and on that basis seeks to clarify the notion of translation/meaning. Interestingly, it is only later, when he republishes his more than ten-year-old papers in a new collection, that he becomes self-aware of his own philosophical agenda:

One thing that only gradually dawned on me was that while Tarski intended to analyse the concept of truth by appealing (in Convention T) to the concept of meaning (in the guise of sameness of meaning, or translation), I have the reverse in mind. I considered truth to be the central primitive concept, and hoped, by detailing truth's structure, to get at meaning. (1984: XIV)

I propose to call a theory a theory of meaning for a natural language L if it is such that (a) knowledge of the theory suffices for understanding the utterances of speakers of L and (b) the theory can give empirical application by appeal to evidence described *without using linguistic concepts specific to the sentences and words of L.* (1984: 215, emphasis added)

Before proceeding, I have to admit my full agreement with Davidson at least as far as one fundamental point is concerned. Like him, I also think that there is nothing more primitive and trivial than the concept of truth. As important as the concept of truth may be, it is easily couched in Tarski's theorems. Truth is what disquotation or Tarski's theorems tell us. The different theories of truth that we find in tradition (adequacy, coherence, and consensus) take us nowhere beyond what is already captured by Tarski's theorems. "The Hard problem" of the philosophy of language, to use the jargon of fashion, is a theory of meaning rather than a theory of truth.

The idea that the truth of the theorems suffices to capture the meaning of the expressions of a language has met much resistance. In the editors' introduction to TM (TM henceforth), Gareth Evans and John McDowell echo the contributions to that volume by J. A. Foster and Brian Loar:

The e fact that each axiom of a truth theory has its impact upon an infinite number of T-sentences does indeed have the consequence that it is difficult for counterfeit theories to pass the test provided by [convention T]. But . . . it is not impossible. Axioms for individual expressions may be chosen which, even though they disperse their inaccuracy, if construed as giving the meanings of those expressions, over as many T-sentences as there are sentences in which the expressions occur, nevertheless preserve the truth of all the T-sentences . . . *It is thus obvious that more stringent conditions must be imposed upon a theory of truth*, if it is to be seen as a theory of meaning, than that its T-sentences be true. (Evans and McDowell, 1976: xiv, emphasis added)

Ramberg rephrases the same objection decades later:

It is possible, given any theory of truth, to derive from it, with the aid of simple logical devices, a second theory that extensionally matches the first, yet would be unacceptable as a theory of meaning. (1991: 61)

Fodor & Lepore bring the same problem to the fore:

So the situation seems to be this: Davidson says at one point that his program is to "take truth as basic and to extract an account of translation or interpretation" ("Radical interpretation," p. 134). The obvious prima facie objection to this project is that, whereas notions like *means that* and the like are intensional, the truth of a T-sentence requires only equivalence of truth value (extensional equivalence) between the formula mentioned on the left and the formula used on the right. lt is thus reasonable to wonder how a theory constrained only to issue in true T-sentences could hope to reconstruct the semantic relations. Clearly, some *further* constraint must be placed on the truth theories that are to count as successful meaning theories (1992: 61-62)

It is always possible to derive from any given theory of truth, with the aid of logical devices, a second theory that matches the first, yet would be intuitively unacceptable as a theory of meaning. Following Fodor & Lepore, let me call this *the extensionality problem*. However, in opposition to Evans, McDowell, Fodor and Lepore, I am firmly convinced that imposing more stringent conditions upon a theory of truth brings no solution to the problem. A successful reductionist theory of meaning has to give up the old shabby analytical dogma that which prioritizes language

over everything and hence to give up the entire framework of radical interpretation.

As usual, Davidson never makes this problem explicit. Still, there is no doubt that he was quite aware of it. It appears tacitly implied in pages 25-26 of his TM:

> Still, this fact ought not to con us into thinking that a theory any more correct that entails '"Snow is white" is true if and only if snow is white' than one that entails instead: (S) 'Snow is white if and only if grass is green'. (1984: 25-26)

In any case, the problem is easily formulated in the following terms. Since every true sentence is materially equivalent to any other equally true sentence, nothing prevents us from constructing theorems such as:

'Der Schnee ist weiss' is true in German *iff* grass is green.

'Der Schnee ist weiss' is true in German *iff* snow is white.

According to Davidson's own words:

> The grotesqueness of (S) is in itself nothing against a theory of which it is a consequence, provided the theory gives the correct results for every sentence (on the basis of structure, there being no other way). It is not easy to see how (S) could be party to such an enterprise, but if it were - if, that is, (S) followed from a characterization of the predicate 'is true' that led to the invariable pairing of truths with truths and falsehoods with falsehoods - then there would not, I think, be anything essential to the idea of meaning that remained to be captured. (1984: 26)

In both cases, the object language is German while the metalanguage is English. However, while intuitively theorem T correctly translates German sentences, S does not do, even though it is implicated by a theory of truth for German in the same way as T is. The extensionality problem can be expressed in the following terms: how is an extensionally adequate theory of the truth-predicate for German able to exclude theorems such as S?

Let me briefly consider two bad suggestions as solutions to the problem. First, Evans & McDowell claim that if a theory of truth for language is to serve as a theory of meaning, it must

be constrained by conditions the formulation of which "would involve employing psychological concepts; and some may see this as raising hopes for the reduction of the concept of meaning to the concept of psychological concepts" (1976: xv). Yet, as soon as we realize that psychological concepts like the concept of belief and the concept of concept are also intensional notions on the same footing as the concept of meaning, we also realize that Evans & McDowell's suggestion is helpless. What we need is a non-linguistic theory of meaning and belief.

Second, Ramberg claims that:

> We should recall that a natural language is never a complete, clearly delineated entity, and hence give up the idea that a language can ever be modeled by a complete truth-theory. For it is only on the assumption that there is some one theory or other that definitively captures the meaning of the words of a language, that we can imagine a situation where convention T could not be used to rule out a 'counterfeit theory'. (1991: 61)

In those words, Ramberg wants us to believe that a theory of meaning must coexist with Theorems like S just because we can never achieve a *complete* theory of truth for a given language. Well, as Davidson himself holds, theorems like S are nothing but grotesque. If a theory of meaning has to live with such grotesqueness as S, so much the worse for it.

Meaning Holism

In his own unsystematic and confused way of putting things, Davidson presents at least three conditions that a theory of truth à la Tarski should satisfy in order to solve the extensionality problem. In TM, he initially suggests that meaning holism together with the compositionality of natural languages could provide a solution to the extensionality problem. Yet, in the same article, he also mentions the putative nomological character of Tarski's theorems of truth as an additional solution. The same suggestion reappears in his article "Reply to Foster". Finally, in "Radical Interpretation," the paper in which Davidson presents the definitive version of his reductionist theory of meaning, he suggests in addition that the principle of interpretative charity could provide a solution to the extensionality problem. Now, what is in fact disconcerting to any Davidson reader is that at no time did he make clear if each new condition replaces the previous

one, or if they should be seen as complementary, or worst of all, if they amount to being the same requirement.

Let me appreciate meaning holism first.[1] Davidson's argument is condensed into a footnote:

> Critics have often failed to notice the essential proviso mentioned in this paragraph. The point is that S could not belong to any reasonably simple theory that also gave the right truth conditions for "That is snow" and "That is white" (1984: 26, n.10)

If a theory of truth for German implies S and respects the compositional structure, then it will attribute erroneous truth conditions to other sentences in German containing the expressions "Schnee" and "weiss", in particular to demonstrative sentences such as "Dies ist Schnee"; "Dies ist weiss". For example, this would mean accepting the theorem:

---

1   A few elementary observations are required. First, generally stated, the principle of compositionality applies not only to complete sentences or propositions whose semantic value is the truth-value, but also to any syntactically complex linguistic expression. It says only that the semantic value of a syntactically complex expression is the value of the function whose arguments are the constituent terms of the complex expression plus the logical syntax of the complex expression.

Second observation: compositionality rests on two trivial empirical facts. The first, known since Aristotle, is the so-called "productivity" of natural languages. The semantic value of a syntactically complex expression is the value of the function that has as its arguments the value of the simple terms constituting the complex expression plus the logical syntax of the complex expression. This is the empirical assumption that best explains our semantic capacity to construct an indefinite number of grammatically and semantically well-behaved sentences, based on a finite lexicon and a logical syntax.

The second decisive empirical factor is systematicity. One who is able to utter "Peter loves Mary" is equally able to think" "Mary loves Peter," though there is no implication between the sentences; after all there are unrequited loves. That the semantic value of a syntactically complex expression is the value of the function that has as arguments the value of the simple terms constituting the complex expression plus the logical syntax of the complex expression is by far the assumption that best explains systematicity as a capacity of any speaker.

Third observation: compositionality is neutral in the face of atomism (building blocks) and of meaning holism. If I adopt a building-block theory, I could first assign semantic value to the component parts of complex linguistic expression and based on its logical syntax assign a semantic value to the complex expression. On the other hand, if I adopt a holistic theory, I would first have to attribute semantic value to sentences of a theory only then to assign semantic values to the components of such sentences. It is this second path that Davidson takes: meaning holism.

'Dies ist weiss' is true in German *iff* this is green.

The first question that arises is whether holism is indeed a necessary condition for eliminating theorems like S. On closer inspection though, what is doing the job of excluding theorems like U is not meaning holism or the compositional nature of German, but rather the empirical data available to the interpreter. When he attempts to confirm U, by saying 'dies ist weiss' pointing to anything green, the speaker will refuse to assent.

Let me take stock. I am firmly convinced that any theory of language that does not satisfy the condition of compositionality is doomed to failure. Nevertheless, as Davidson invariably speaks of compositionality as a *constitutive condition*, he seems to suggest that there could be no natural language that did not satisfy the principle of compositionality. I wonder: could we rule out *a priori* a language of unstructured expressions? I do not think so. Wittgenstein provides a counterexample:

> Let us imagine a language for which the description given by Augustine is right: the language is meant to serve for communication between a builder A and an assistant B. A is building with building stones: there are blocks, pillars, slabs and beams. B has to pass him the stones and to do so in the order in which A needs them. For this purpose they make use of a language consisting of the words "block", "pillar", "slab", "beam". A calls them out; B brings the stone which he has learnt to bring at such-and-such a call. —— *Conceive of this as a complete primitive language.* (2009: § 2: 6, emphasis added)

Here goes another counterexample. Let us suppose two individuals (1 and 2), identical in many aspects, but distinct in a relevant mental aspect: while the former would have a standard memory of any person of his age and education, the latter has his brain wirelessly connected to a supercomputer from which his memory is fed by an impressive amount of data. While individual 1 states "snow is white" to say snow is white, subject 2 makes use of the unstructured expression "B" to mean the same. Likewise, while individual 1 uses the sentence "this is snow" to say this is snow, individual 2 uses the unstructured expression "N" to mean the same. And when individual 1 uses the sentence "this is cold" to say this is cold, individual 2 uses the unstructured expression "F" to means the same, and so on. Now, we may suppose that when circumstances induce individual 1 to utter: "this is snow," the same circumstances induce individual 2 to utter

"N." If the first individual is able to infer from the sentence "this is snow" the sentence "this is cold," the second is able to infer "F" from "N," and so on. I wonder: what is the argument that proves that the language of 2 is impossible *a priori*?

Here is a methodological aspect of Davidson's philosophy (not only his philosophy of language but also his philosophy of mind) that is often neglected: his shamefaced transcendentalism. If on the one hand Davidson declares himself a naturalist, on the other, he erects its principles (compositionality, interpretive charity, etc.) into the condition of *constitutive* principles, that is, *a priori* conditions of understanding. There is a tension here that every Davidson interpreter has to solve between Davidson's semantic transcendentalism and Davidson's Naturalism.

But if meaning holism plus compositionality does not seem to be a *necessary condition* for excluding from our theory of truth German theorems such as S, the question is whether it would be at least a *sufficient condition* for the same ends. The answer is clearly negative when we consider that the predicates involved can be coextensive without being synonyms such as triangular and trilateral predicates. How to exclude theorems from form:

'A is triangular' is true in English *iff* A has three sides.

Davidson's Principle of Charity

Let us now turn our attention to the principle of interpretive charity. Precise articulation of the principle of charity turns out to be a tricky task when we take into account Davidson's reductionist semantic. We find in Davidson's work at least three different formulations. The first is the normative principle that we should maximize the truth of most of a native's utterances in order to make sense of his utterances. The second is the normative principle that we should maximize agreement between the speaker and us interpreters in order to make sense of his utterances. And the third is the normative principle that we should maximize the consistency in the native's system of beliefs in order to make sense of his utterances. The fundamental idea is that the interpreter can only find the speaker's utterances false, disagree with the speaker, or find his beliefs inconsistent, after he has ensured the understanding of most of the speaker's utterances.

On closer inspection though, in light of the extensionality problem and Davidson's own reductionist program, only the first formulation, viz. maximization of the truth of the native's utterance, is acceptable. Unfortunately, Davidson usually confounds the first two and more seldom the three formulations and by doing so he obfuscates the significance of the principle and of his project[2]. Consider this:

> The method is intended to solve the problem of interdependence of belief and meaning by holding belief constant as far as possible while solving for meaning. This is accomplished by assigning truth conditions to alien sentences that *make native speakers right* when plausibly possible according, of course, to our own view of what is right. What justifies the procedure is the fact that *disagreement and agreement* alike are intelligible only against a background of massive agreement. Applied to language this principle reads: the more sentences we conspire to accept or reject (whether or not through a medium of interpretation), the better we understand the rest, whether or not we agree with them. (1984: 137, emphasis added).

> The basic methodological precept is, therefore, that a good theory of interpretation maximizes agreement. Or given that sentences are infinite in number, and given further considerations to come, a better word might be optimized. (1984: 169)

> The method is not designed to eliminate disagreement; nor can it. Its purpose is to make meaningful disagreement possible and this depends entirely on a foundation- some foundation- in agreement. (1984: 196-197)

> Charity is forced on us; whether we like it or not, if we want to understand others, we must count them right in most matters. (1984: 197)

The problem is that the concept of belief (and therefore also the concept of agreement) is an intentional concept. And its job is precisely to show how we get the intentional from the extensional. Belief and meaning must arise in the same way, simultaneously, by the imposition of structure on an infinite supply of extensional evidence. Yet, as a methodological principle, charity must apply also in the initial stages of theory construction. And since at this point there is nothing to which the interpreter can match her beliefs, no notion of matching beliefs or of agreement will capture the task of interpretation.

---

2   See 1984: 137, 153, 169, 196-7, and also Davidson, 1980: 221, 290.

That is the reason why we cannot take seriously Luke's suggestion of replacing Davidson's charity by Richard Grandy's principle of humanity:

> The Principle of Charity counseled 'Count them right in most matters'. The Principle of Humanity counsels 'Count them intelligible or perhaps count them right unless we can't explain their being right or can better explain their being wrong'. In other words, it prescribes the minimizing of unintelligibility - that is, of unintelligible agreement and disagreement. It has the singular virtue of being the principle we do in practice apply on the interpretation and translation of beliefs. (Lukes 1982: 262)

Lukes holds that Davidson's unqualified charity bases the necessary agreement on too many truths. He implicitly construes the principle as one to be applied in choosing between alternative existing translation schemes, because it supposedly works on the attribution of beliefs, something which is not possible until some kind of a theory of meaning is in place.

However, in at least one passage, Davidson seems to suspect that the principle of charity formulated in terms of agreement is unacceptable:

> My point has always been that understanding can be secured only by interpreting in a way that makes for the right sort of agreement. The "right sort", however, is no easier to specify than to say what constitutes a good reason for holding a particular belief' (1984: xvii).

In short, considering Davidson's writings, the principle of charity can be interpreted both in an extensional and in an intensional way. In a purely extensional way it reads that we must maximize the truth of what the speaker says to understand what is said. The problem faced by the interpreter is that of establishing a connection between the sentences of L and the observable circumstances of their being uttered. The extensional link is, of course, truth. Read, however, in the intensional form, the principle reads that we interpreters must maximize our agreement with the speakers by attributing our own beliefs to them. It goes without saying that in the intensional reading of the principle of charity Davidson begs the question at issue: in order to solve the extensionality problem he would be consciously or unconsciously giving up his project of formulating a theory of meaning on purely extension bases. He would exclude the theorem:

'A is cordate is true in English *iff* A has kidneys.

But he would do it because in order to maximize his agreement with the speaker he would have to assign his own belief to the speaker, in that he would have to give to "cordate" the same meaning it has in his language. In this way Davidson would be violating at the same time the restrictive conditions that he and Quine impose on a radical interpretation.

The alternative would be to read the principle of interpretive charity in an exclusively extensional way in accordance with the constraints of radical interpretation. In these terms, however, the appeal to the principle of charity provides no solution to the extensionality problem. If the cordate and renate predicates are coextensive, we will be charitable in interpreting their utterances by both Theorems W and:

A is a cordate is true in English iff A has a heart.

Lawlike Connections

Now let us turn to the requirement that theorems of theory must be laws. Again, for absolute lack of clarity, Davidson only clarifies how the nomological statute of the theorems à la Tarski could solve the problem of extensionality in a miserable note in the 1984-edition of TM:

> This paragraph is confused. What it should say is that sentences of the theory are empirical generalizations about speakers and so be true but also lawlike. (S) presumably is not a law since it does not support the appropriate counterfactuals. It is also important that evidence for accepting the (time and speaker relativized) truth conditions for "that is snow" is based on causal connections between a speaker's assent to the sentence and the demonstrative presentation of snow. (1984: 26, n.11)

What Davidson has in mind are generalizations of that form:

(GE) (x)(t) (if x belongs to the German speech community, then x holds true 'es regnet' on Saturday at t *iff* it is raining near x at t)

Therefore, it is false that:

(GE) (x)(t) (if x belongs to the German speech community, then x holds "Die Schnee is weist" on Saturday at t iff grass is green near x at t)

According to Segal, Davidson's strategy is unsuccessful because Since it is a law that electrical and thermal conductivity would coexist in metals, the scientific statement (3) below supports a counterfactual construction:

"Copper conducts heat" is true in English iff copper conducts electricity.[3]

Now, Segal's objection would only be effective if the properties of *conducting heat* and *conducting electricity* were in fact coextensive (like triangular and trilateral), but they are not: there are cold conductors of electricity, the so-called superconductors. Given this, Davidson could retort that we must add the requirement that sentences of the theory follow laws of meaning holism: "this is copper"; "This leads to heat"; "That drives electricity". To the extent that the speaker could assent to "this conducts electricity," but not to "that conducts heat" (pointing to a superconductor), the interpreter could rule out theorem number Y.

However, the obvious objection is that sentences of the theory cannot be laws because they depend on linguistic conventions that are purely contingent. We use the phrase "snow is white" to say that snow is white thanks to numerous conventions of English. So, if S is not a law because it does not support the appropriate counterfactuals, neither is T: in possible worlds close to the actual world, one could utter "Gavagai" in English to say that snow is white. Moreover, even if such sentences were laws, they *per se* do not discriminate coextensive predicates. Consider again:

'A is triangular' is true in English *iff* A has three sides.

(x)(t) (if x belongs to the English speech community, then x holds 'A is triangular' on Saturday at t *iff* A has three angles near x at t.

If S is a law, then V above is also one. The moral is that there is no nomological solution to the extensionality problem *within a linguistic framework like Quine's radical translation or*

---

3   See Segal 1999.

Davidson's radical interpretations. Indeed, assuming the linguistic turn, that is, the idea that there is priority of language over thought or that they are on the same footing, Quine is closer to the truth than Davidson. The thesis of indeterminacy seems plausible.

However, it is plausible to assume that there are laws here that are not linguistic or formulated in terms of contingent linguistic conventions. For example, it is quite plausible to assume that there are laws here connecting brain states and mundane facts, for example, between some brain states of someone contemplating the snow and the fact that snow is white. It is reasonable to suppose that in worlds close to ours when creatures like us are faced with white snow, their brains are all in the same global state.

(x)(y)(t) (if x & y belong to the human race, then x & y are in the same brain state on Saturday at t *iff* they see snow near them at t.

It goes without saying that this lawlike connection between some brain states and the fact that snow is white is not enough to ensure that the brain *represent* the mundane fact that snow is white. Much more is needed. The first crucial assumption is that a particular brain state that nomologically covariates with the color white is "recruited" due to evolutionary reasons to indicate that color so that when that state fulfills that indication-function it veridically represents the color white; otherwise it has misrepresented it.

Something similar occurs with the brain state that nomologically covariates with the presence of snow. *Per se*, that brain state does not represent the snow. However, that it may be recruited due to evolutionary reasons to indicate the presence of snow so that when the state fulfills that indication-function it veridically represents the snow; otherwise it has misrepresented it. Now what is worth noticing is that the "recruitment" is the crucial moment when the brain state becomes a mental state endowed with a natural meaning, namely to indicate the presence of white. I do not appeal to any intensional notions like Davidson. Instead, I account for it in terms of something that is neither intensional not mental: biological functions.

Now, as the German sentence of the object language 'Der Schnee ist weiss' expresses a *belief*, something more is needed. We must assume in addition that SNOW and WHITE are concepts in the speaker's mind. Thus, by learning processes those brain states are "recruited" a second time to indicate not a particular token of white or a particular token of snow, but rather every type of snow and white for which those particulars are tokens.

Given this, we may rule out S as a bad translation of the German sentence of the object language 'Der Schnee ist weiss' because S does not express a lawlike connection and because the speaker's brain states when he utters 'Der Schnee ist weiss' do not have the function of indicating the fact that grass is green. Now, in this picture Davidson's famous triangulation assumes another form. To be sure, people share the same environment. Moreover, that environment elicits in their brains the same states and in virtue of evolution and learning they also share the same belief, namely that snow is white.

My proposal nicely handles the cases of coextensive predicates with different meanings. Let us take a closer look:

'The dog has a heart is true in English *iff* the dog is a renate.

The dog has a heart is true in English iff the dog is a cordate.

The nomological covariation between brain states and mundane fact is purely extensional and hence cannot rule Y out: if it is a law that all cordates have a heart, it is also a law that all cordates have kidneys. Likewise, the states of brain that covariate with renate animals also covariate with cordate animals. Still, if the nomological connection dos not cut the ice, the teleological element can easily capture the different meaning of coextensive predicates. Why so? Because between all brain states that covariate with cordates and renates, the learning process "recruits" a few of them to indicate cordates while the learning process recruits others to indicate renates. That is enough to rule out Z as a bad translation and to endorse AA.

I would like to finish this paper with some methodological observations. First, Davidson is unable to find a solution to the extensionality problem because he remains imprisoned by the linguistic turn. To be sure, he disagrees with Dummett when the latter claims the priority of language over thought, the key dogma of analytical philosophy. Yet, he still believes that language and thought are at the same level of analysis when clearly thinking precedes language. To learn a language and to translate a language is nothing more than to learn how to express previously existent thoughts and beliefs.

Now, a Davidsonian may raise the following key objection. Quine and Davidson's constraints on radical interpretation make nothing available as empirical data apart from the speaker's linguistic and non-linguistic behavior. The field linguist has the job of translating a

ROBERTO HORÁCIO DE SÁ PEREIRA

completely unknown language that has no historical or cultural connections with any known language, and for which the linguistic has no manuals or recourse to bilinguals. Given this, how does the interpreter figure out in which of the two brains states (one having the function of indicating renates while the other the function of indicating cordates) the speaker is in when a rabbit crosses his way and he utters: "Gavagai"? Indeed, there is no way. But and so what? That only shows that there is something deeply wrong with Quine and Davidson's framework. A theory of meaning cannot be provided as a radical interpretation.

RESUMO

O projeto semântico de Davidson toma a forma de reducionismo que visa explicar as noções intencionais com base em noções puramente extensionais. O objetivo desse trabalho consiste na investigação do fracasso do projeto de Davidson como um argumento indireto contra sua suposição segundo a qual a chamada interpretação radical seria o fundamento do significado linguístico e do pensamento.

Palavras-chave *Davidson; interpretação radical, o problema extensional, reducionismo semântico.*

ABSTRACT

Davidson's semantic program is a form of semantic reductionism that aims to account for intensional notions on a purely extensional basis. My aim in this paper is not to show that Davidsons fails in addressing the so-called extensionality problem. That is no novelty. Even Davidsonians recognize Davidson's failure. My main goal is to explore Davidson's failure as an indirect argument against his key assumption that the radical interpretation is the basis of sense and hence the possibility of thought. There is no priority of language over thought. This is a remnant of the so-called linguistic turn that Davidson inherited from the analytical tradition. The moral is the following: we must come back to Lockean common sense that to learn a natural language and to translate an alien tongue into our own is merely to learn a vehicle for communicating and expressing thoughts we have already had.

Key-words *Davidson; Radical Interpretation; extensionality problem; Semantic Reductionism*

References

Davidson, D. (1980). *Essays on Actions and Events.* Clarendon Press, Oxford.

Davidson, D. (1967/1984). "Truth and Meaning". In *Inquiries into Truth and Interpretations*. Oxford University Press, New York, 17-36.

Davidson, D. (1973/1984). "Radical Interpretation". In *Inquiries into Truth and Interpretations*. Oxford University Press, New York, 125-40.

Davidson, D. (1976/1984). Reply to Foster". In *Inquiries into Truth and Interpretations*. Oxford University Press, New York, 171-79.

Lepore. E. & Ludwig K (2002). *What is a Logical Form. In Logical Form and Language.* Preyer. G & Peter G (eds), 54-90. Oxford University Press, Oxford.

Quine, W. V. O. (1951)."Two Dogmas of Empiricism. *The Philosophical review* 60 (1): 20–43.

Quine, W. V. O. (1960). *Word and Object.* The MIT Press, Cambridge, Massachusetts.

Segal, G. (1999)."How a Truth Theory Can Do Duty as a Theory of Meaning". In Donald Davidson: Truth, Meaning and Knowledge. Urszula M.Ž eglen´ (ed.), Routledge, London.