

# DAVIDSON ON COMMUNICATION AND LANGUAGES: A REEXAMINATION

---

**FELIPE CUERVO**

*Universidad El Bosque*  
*Departamento de Humanidades-Filosofía*  
*Av. Cra. 9A #131A-02 Bogotá 110121*  
*Colombia*  
fcuervor@unbosque.edu.co

## **Article info**

CDD: 401

*Received: 22.02.2018; Revised: 13.09.2018; Accepted: 10.10.2018*

DOI: <http://dx.doi.org/10.1590/0100-6045.2018.V41N3.FC>

## **Keywords:**

Conventions

Radical interpretation

Donald Davidson

**Abstract:** In order to evaluate the validity and implications of Donald Davidson's arguments against the need for conventions in order for linguistic communication (or, more generally, against the need to postulate language as an entity in order to account for communication), the theoretical considerations behind his conclusions are traced through several of his essays. Once Davidson's ideas on communication, radical interpretation, and the lack of strict nomological connections between physical and mental events have been pointed out as necessary for his argument, it will be seen that these imply the need for something very close to linguistic conventions. The article closes by considering a few possible counterarguments this last conclusion.

If we were to characterize Donald Davidson's philosophical work, "controversial" would surely be one of the first words that comes to mind, and among the philosophical disputes he initiated, perhaps none has provoked such heated debate as his rejection of language, understood as previously learned rules or conventions (Davidson 2005:110) as relevant for a philosophical understanding of linguistic communication. Replies have included rather timid acceptance of his conclusions as describing communication that happens not to be human (Lepore & Ludwig 2005: 296-297), accusations of having misused the notion of interpretation (Hacking 1986: 449-451), and rejection of the explanatory utility of his ideas (Green 2001: 252-253; Reimer 2004: 331-332; Stainton 2016: 18). Rather than assume an external position from which to evaluate Davidson's ideas, the purpose of this essay will be to consider them in the light of his other theories concerning communication, hoping thereby not only to use his own ideas to defend the importance of conventions, but also to offer some clarification of what the role of conventions in language is supposed to be.<sup>1</sup>

This might seem a rather ambiguous position, so, before setting out, I want to offer a brief characterization of where I will be standing; I believe (and will attempt to demonstrate) Davidson's conclusions regarding conventionality cannot be defended, even within his own system. If I can demonstrate that the model of linguistic

---

<sup>1</sup> I wish to thank the two anonymous reviewers for *Manuscrito*, who have, besides pointing out several obscurities in my arguments and offering insightful counterarguments, recommended several texts dealing with the same issues as this paper of which I was not aware.

communication interpreted by Davidson as denying the need for linguistic conventions relies on his previous theoretical work on meaning and that this same work implies the need for conventions in language, then I will have proven that either the model is mistaken or Davidson's interpretation of what it implies does not follow from it. I believe the second to be the case, but, given evaluating the model as a whole would require a different investigation, will be content to offer a new interpretation that defends the need for conventions, hoping that those who, as I do, believe Davidson's description of linguistic interaction is both essentially correct and fruitful, but are not willing to follow him all the way to the denial of the conventionality of language, can make use of his ideas without a bad conscience.

Section I offers a schematic account of his argument against language and the general theory of linguistic communication on which it depends; section II presents an overview of radical interpretation, emphasizing those aspects relevant for section I's argument; section III, after a small investigation into the nature of arbitrariness, examines the coherence of Davidson's arguments against language and conventions with the theoretical apparatus discussed in section II, and section IV attempts to answer three possible counterarguments to section III's conclusion.

## I

Davidson's argument against language (understood, as mentioned above, as a set of rules or conventions that must be learned before linguistic communication succeeds)

consists of two parts:<sup>2</sup> 1) proving language is not sufficient for linguistic communication, and 2) proving language is not necessary for linguistic communication (Davidson 2005: 110).<sup>3</sup> We will approach each branch of the argument separately and afterwards introduce a couple of general considerations.

1) In order to prove language is not sufficient for linguistic communication, Davidson relies on his well-known analysis of successfully communicative malapropisms. To use his own example, when Mrs. Malaprop (or, perhaps, Sheridan) utters “a nice derangement of epitaphs”, we all understand her as meaning “a nice arrangement of epithets”, and thus can make sense of the totality of her message. Given there is no prior convention according to which, for example, “epitaph” might mean in certain contexts “epithets” (in other words, no standard dictionary would include “epithet” under its entry for “epitaph”), knowledge of conventional meaning cannot provide an explanation for the fact the we understand Mrs. Malaprop. We must rely on non-conventional, possibly non-linguistic, sources in order to account for communicative success. The conclusion is quite evident: if communication succeeds and its success cannot be explained by conventions, then knowledge of conventions cannot be sufficient for communication (Davidson 1986: 90-91).

---

<sup>2</sup> Both Stainton (2016: 9-11) and Lepore and Ludwig (2005: 269-71, 278) offer a detailed analysis of his argument in a similar manner.

<sup>3</sup> Since the evolution of Davidson’s ideas will play an important part in my analysis, I quote his essays according to the year they were published.

It must be kept in mind that objections such as those raised by both Green (2001: 250-251) and Cain (2013: 147-148), according to which malapropisms are understood only against the background of conventional or standard meaning (we know Mrs. Malaprop means “epithets” because we recognize how similar its sound-pattern is to that of “epitaph” and we know the former word exists in English) cannot be directed towards this branch of the argument: they prove it is necessary for us to have knowledge of language in order to account for malapropisms, but such knowledge might still be insufficient for successful communication. If we are unaware, for example, of the context in which a malapropism is uttered, we might be unable to identify it as a malapropism, and thus would have no reason to attribute a non-standard meaning to her words. In such a case, we would not grasp the speaker’s intended meaning and communication would have failed.

2) Davidson’s argument against the necessity of language depends on the validity of his general description of linguistic communication, according to which “what must be shared is the interpreter’s and the speaker’s understanding of the speaker’s words” (Davidson 1986: 96). In order for this to occur, even if the interpreter starts out with a prior theory<sup>4</sup> of what certain words mean, and the speaker starts out with a prior theory of what the interpreter’s prior theory might be, communication will not succeed unless the way the speaker actually intends her

---

<sup>4</sup> When used in the context of prior and passing theories, “theory” refers not to an actual theoretical construction, or even conscious knowledge, on the speaker’s or interpreter’s part, but to the description an observer must develop in order to explain successful communication between them.

words to be understood (the speaker's passing theory) and the meaning her interpreter attributes to them (the interpreter's passing theory) converge (Davidson 1986: 100-101). If this account is correct, then only the convergence of passing theories is necessary for successful communication, and we must then ask ourselves whether this necessity somehow implies that language is necessary as well.

Prior theories might initially seem to be good candidates for language, specially considering they comply with the requirement that language be prior to actual linguistic communication, but a quick glance at the schematic description of communication offered in the previous paragraph shows there is no apparent need for prior theories to match (*cf.* Davidson 1986: 103). As with malapropisms, we might discover that appealing to the usual way we understand certain words will make a speaker's utterances seem contradictory, incoherent, or quite simply unintelligible, while relying on non-standard meanings makes her words perfectly sensible. If we manage to find the necessary non-standard meanings, and they coincide with those the speaker intended her words to have, our passing theories will converge and communication will succeed without the need for matching prior theories. No such a thing as shared conventions would be necessary for communication.<sup>5</sup>

---

<sup>5</sup> In "Communication and Convention", Davidson explicitly directs a first version of the arguments presented above at David Lewis' well known theory of conventionality, under the assumption that Lewis' conventions must have been decided upon prior to the encounter between hearer and speaker (Davidson: 1982). Josh Armstrong has defended Lewis, claiming Lewis' original theory was meant as a solution to a problem raised by Rousseau, Russell, and Quine (communication cannot depend on conventions, since we need to communicate in order to decide

It might be argued against this that, even if prior theories diverge slightly, they must have something similar to a common nucleus against which mistakes are identified and, if possible, interpreted (this would be the proper place to consider the objections mentioned in (1)). This would forbid the use of malapropisms and similar mistakes as arguments unless we find a way of demonstrating that their interpretation does not rely on knowledge of conventions, and would force Davidson to prove communication is still possible even if prior theories differ completely; in other words, Davidson would have to demonstrate communication is possible even when we can rely on no knowledge whatsoever of the speaker's language or anything that we could know about her only through language. Since these are the conditions under which radical interpretation is supposed to take place and that is a subject we will discuss in section II, the solution to this problem will have to be postponed for the moment.

The problem we discussed a moment ago brings us back to our other candidate: passing theories. If communication can succeed when prior theories are radically different, it is only because passing theories can be made to converge during communication. So it would seem that, if something

---

upon conventions), and so included an account of how conventions could be established without relying on prior conventions (Armstrong 2015: 86-88, 95-97). Armstrong has also provided a dynamic account of conventionality that can deal with shifts in lexical meaning (Armstrong 2016: 100-110). Even though I believe there is much to be gained from studying Lewis' work and will retake some of the arguments Armstrong offers on his behalf in the last section of this essay, a detailed exegesis of Lewis' work would lead us too far from our current goal, which is to examine the internal coherence of Davidson's argument. I will thus rely on, mistaken as it may be, Davidson's portrayal of Lewis' ideas in order to grasp Davidson's arguments.

can be called a language, it would have to be a passing theory. But passing theories need not be permanent; that is to say, as long as a mechanism that makes the convergence of passing theories possible exists, there is no need for a passing theory to last more than a single encounter between a speaker and an interpreter. Inefficient as it might turn out to be, theoretically speaker and interpreter might be forced to recommence the whole process that leads to convergence every time they interact linguistically, and communication would still succeed, even if slowly and painfully. Calling a passing theory a language would mean that we would have to consider each encounter's passing theory a new language, and would be forced to consider every comprehensible malapropism and mistake, given they imply a change in the interpreter's passing theory, as causing a language change. A single conversation might then occur in a multiplicity of languages, which would make the concept of language seem rather useless (*cf.* Davidson 1986: 102, 106). Even more, assuming mistakes and such are not predictable, and that words can acquire meanings useful only in the very specific circumstances of a determinate linguistic encounter (words such as “thingamajig” seem to have such a use), there would be no way for the interpreter to learn, prior to communication, everything her passing theory would need in order to converge with the speaker's. Languages, then, would be neither learnable nor exist prior to actual communication, which goes against our initial definition of them (*cf.* Davidson 1986: 105-106).<sup>6</sup>

---

<sup>6</sup> Peter Ludlow (2014) has recently developed a theory of meaning as not merely lacking clear edges, but as essentially underdetermined, lacking anything similar to a nucleus around which doubtful applications might be gathered, and which receives further determinations (which, of course, never



If my interpretation of Davidson's argument is correct, then the necessity of language for communication can only be proven false if he can provide us with a mechanism that permits passing theories to converge even when there is no prior shared element. In other words, unless we can demonstrate it is possible to interpret a speaker of whose language we have no knowledge, we have no reason to accept the second branch of the argument. But what about the first? Apparently, even if such mechanism does not exist, his argument holds, but, by itself, it does not seem to prove anything of much interest. As several of his critics have stated, most conventionalists are willing to accept language is necessary but not sufficient for proper interpretation,<sup>7</sup> and if we are left with only this half of his

---

completely determine the meaning of any word: being further determined never implies ceasing to be underdetermined) according to local needs: these further determinations can become the subject of either semantic imposition or litigation among speakers. According to Ludlow, whose work is well supported by analyses of how meaning-fixation actually takes places (2014: 39-71), his ideas can be seen as an extension of Davidson's arguments, interpreted as proving languages exist, at most, as microlanguages that apply only to specific linguistic encounters (2014: 96-97). In general, I agree with Ludlow's analysis, but believe they can be accounted for within a conventional theory of language (such as Armstrong's dynamic model; *vide* note 4 for references). Without discussing his ideas in detail (I hope to do so in a future essay), all I wish to say here is that meaning, even if it lacks a stable nucleus, must have some kind of recognizable identity in order to justify interpreting and event as a change in meaning instead of the creation of a new lexical item after each dispute (even if this unity is guaranteed by nothing more than Millikan's lineages) (Millikan 2003: 34).

<sup>7</sup> See, for example, Stainton (2016: 11-15), as well as Lepore and Ludwig (2005: 275). Even David Lewis, one of Davidson's main

argument, we could conclude the pragmatics of language are as important as its conventional elements, but would not be able to get rid of the latter in favor of the former, which seems to be Davidson's intention. Consideration of his reasons for rejecting conventions thus leads us to his ideas on radical interpretation.

## II

Radical interpretation is Davidson's solution to two different problems: (a) What can we know that would allow us to determine what a speaker's words mean? (b) What evidence, that does not presuppose knowledge of the speaker's language, can an interpreter rely on in order to support his interpretation? (Davidson 1973a: 125; *cf.* Lepore & Ludwig 2005: 152).<sup>8</sup> Assuming a truth theory for a language from which only the T-sentences for every sentence in the same language can be proven provides us with a theory of meaning for the language in question (*cf.* Davidson 1967), Davidson proposes that the interpreter begin by determining, based on his behavior, which sentences the speaker holds true. Applying the principle of charity, she will assume that most of the sentences the speaker holds true are, in fact, true (at least according to her knowledge); she must then correlate the set of true sentences to what, in the circumstances in which they were

---

targets in *Communication and Conventions* (1982), accepts that not every aspect of language is conventional (Lewis 1983: 180).

<sup>8</sup> A much more detailed presentation of the process of radical interpretation than that which follows, which has greatly increased my understanding of Davidson's work, can be found in Lepore & Ludwig (2005: 196-7).

uttered, she would both hold true and consider relevant, thus producing a T-sentence for each original sentence (“for speaker S, at time t, sentence s is true for S at t if and only if p”). Once she manages to formulate a theory that can account for the set of T-sentences, she would be in possession of a hypothetical theory of meaning for the speaker’s language and thus would have answered, at least partially, question (a) (*cf.* Davidson 1973a: 135-138). Nevertheless, she would still be unable to answer question (b); in other words, she would still be incapable of proving her interpretation is correct. It might happen, after all, that what she considers relevant under the circumstances of her encounter with the speaker has nothing to do with what the speaker considers relevant, which would mean the interpreter might have ended up by attributing to the speaker a set of beliefs that are unrelated to the meaning she intended to communicate (*cf.* Davidson 1976: 173).

In order to answer question (b), the interpreter must appeal to empirical verification; that is to say, she must “test the theory by sampling its [empirical] implications for truth” (Davidson 1974a: 142). We can imagine several different scenarios that might prove her interpretation is correct: she might be lucky enough to hear the speaker utter the same sentence at a different time, which would allow her to verify whether the circumstances under which she assumed the sentence proved true still applied. She might also try uttering the sentence and, depending on how the speaker reacts, examine what has changed in the circumstances of its utterance that might explain a change in the truth-value attributed to it, or what has remained constant in order to explain why said truth-value was not altered. Perhaps other experiments can be thought up; the important point here is that they all depend on a correlation between what a sentence means and the context in which it is uttered. Since, according to Davidson, the interpreter can

access a speaker's intentions only through language (Davidson 1973a: 127), if no such correlation were to be found, the interpreter would have no other empirical evidence with which to verify the validity of her interpretation. In other words, she would never know if the way she interprets a speaker is even vaguely close to being correct.<sup>9</sup>

The empirical verification processes mentioned above depend on the bi-conditional structure of T-sentences, since, by contraposition, each one of them might be transformed into “for speaker S, at time t, if not-p, then sentence s is not true for S at t”, which is what allows us to reject p as the meaning of s in case p is still true when p no longer is. We might restate this by saying that T-sentences have a law-like behavior that can be tested counterfactually.<sup>10</sup> Since such counterfactual tests seem to be the only method an interpreter has to reject false

---

<sup>9</sup> The need for complex empirical verification in order to reach a valid interpretation is used by Goldberg (2004) in order to criticize radical interpretation because of its incapacity to explain testimonial knowledge; the quite specific approach this essay assumes does not allow me to discuss his ideas in detail, but his discussion of radical interpretation in terms of warranted interpretation greatly influenced my own ideas concerning Davidson's work.

<sup>10</sup> Davidson describes a successful theory as “one that in fact can be projected to unobserved and counterfactual cases” (1976: 174), and T-sentences as “laws which state the truth conditions not only of actual utterances but also of unspoken sentences” (1999: 688). Lepore and Ludwig, discussing the same passages, describe Davidson's objective as “*a true counterfactual supporting truth theory* [emphasis in the original]”, and offer a careful discussion of why the law-like behavior of T-sentences is necessary for empirical verification (2005: 161-163, specially note 132).

interpretations (assuming, of course, she has already applied the principle of charity and made sure the beliefs attributed to the speaker are coherent), the law-like character of T-sentences is what ensures that radical interpretation answers question (b) and thus is necessary for Davidson's project. We will make use of this conclusion in section III, but before that there is one last issue concerning radical interpretation we must discuss.

Judging by his earlier essays, one of the issues that initially led Davidson to his truth theoretical approach to meaning and the possibility of radical interpretation was his insistence on theories of meaning being capable of explaining the learnability of natural languages. Humans are capable of producing and understanding infinite sentences within any given natural language, but, given humans are also mortal and, perhaps, epistemologically limited, they cannot learn the aforementioned infinity of sentences one by one; this apparent paradox requires an explanation, which, according to Davidson, can only be provided by assuming natural languages have a finite number of semantical primitives from which infinite sentences can be constructed recursively (Davidson 1964: 8-9). Thus, words make a "systematic contribution to the meaning of sentences in which they occur" (Davidson 1967: 22); in other words, Davidson requires any valid theory of meaning to be compositional. Given the same requirement is still present in his essays *Radical Interpretation* (1973a: 127) and *Belief and the Basis of Meaning* (1974a: 151), radical interpretation must provide us with a method not only of determining the meaning of sentences, but also of the role their constituents play in the meaning of a sentence as a whole. This requires a slight modification of the experiments mentioned a couple of paragraphs ago, since the interpreter would have to recognize the presence of one same constituent in several sentences and then, for example,

when confronted with a sentence that differs from one that contains the constituent only because the constituent is absent or has been replaced by something else, examine whether the relation between the new sentence's truth value and its context of utterance has changed. The process becomes much more complicated, but, in essence, seems to be the same as that required to determine the correct interpretation of sentences, and thus depends on a similar condition: the role a constituent plays in a sentence must have a law-like behavior that can be subjected to counterfactual tests.

Before moving on, we must confront an issue concerning the principle of compositionality: it is difficult to determine whether this is still a requirement for the general model of communication Davidson endorses in *A Nice Derangement of Epitaphs*, and his statement that the systematicity of language survives “when understood in rather unusual ways” (1986: 107) is, at best, rather obscure. If we add to this that his famous dictum “there is therefore no such thing [as a language] to be learned” (*ibidem*) denies, at least at first sight, the premise for the learnability argument, and, but that Davidson continues to refer to the systematicity of language in essays written after *A Nice Derangement of Epitaphs* (e.g., 1991: 210; 1993: 170),<sup>11</sup> so that textual evidence apparently endorses something along the lines of compositionality, the situation becomes quite frankly confusing. Some critics believe Davidson continues to assume language is systematic (e.g. Hacking 1986: 455-456), while others, such as Green, maintain that giving priority to passing theories and their *ad hoc* attribution of

---

<sup>11</sup> This is especially explicit in his *Reply to James Higginbotham*: “This is the level [that of first meanings] at which a systematic recursive characterization of truth conditions can hope to characterize an aspect of linguistic understanding” (1999: 687).

meanings would force Davidson to relinquish systematicity (Green 2001: 248-249), despite textual evidence to the contrary. Now, whether natural languages are in fact compositional and how those who deny their compositionality account for their learnability lies beyond the scope and intentions of this essay, a rather vague remark will have to suffice for the moment. Pietroski has convincingly argued that Davidson can deny the existence of language without thereby rejecting the existence of conventional first meanings (dictionary meanings, so to say) that behave systematically and so can be described recursively; in fact he maintains that, even though there can be no such a thing as a theory of truth in L valid for every user of a supposed language, there are enough elements in common among various passing theories to justify building partial theories of truth in L that would behave recursively. (Pietroski 1994: 99-100, 107-9).<sup>12</sup> . More generally, given radical interpretation forces us to commence by interpreting the meaning of sentences and only later progress until we reach their constituents, even accepting compositionality no longer applies, which would mean each sentence must be interpreted as a closed unit, still leaves us with the law-like behavior of T-sentences, and that is enough to reach this essay's conclusion. Thus, even though we will occasionally refer to sentential constituents, the

---

<sup>12</sup> According to Pietroski, Davidson's point would be that an account that considers only these elements can never explain actual communication; interpreters, according to him, must also rely on General Intelligence, for which there is no recursive explanation (1994: 104-105). Although Pietroski's interpretations seems to me to be quite on track, I believe his attempted defense fails, given, as mentioned in note 6, defending the conventionality of language does not mean every single aspect of linguistic use must rely on conventions.

argument in section III will be constructed so as to be valid even if radical interpretation must stop at the level of sentences.

### III

If the account of radical interpretation offered in section II is accurate, we may conclude that the interpretation mechanism required by Davidson's arguments against language (section I) has its own requirement: the law-like behavior of T-sentences. In order to reach our final conclusion, a short examination of what this implies is in order. But, first, we will discuss the notion of arbitrariness, which will prove of great use in our argument.

Davidson begins his criticism of conventions by accepting the arbitrariness of language, but denying that what is arbitrary is necessarily conventional (1982: 265). As far as I can see, he never defined arbitrariness<sup>14</sup> or offered something in the way of a proof, but I believe examining his ideas on the impossibility of strict psychophysical nomological connections can both justify his position and prove useful when attempting to understand radical interpretation, and so we will briefly<sup>15</sup> consider this subject.

---

<sup>14</sup> Given the conjunction of "arbitrary" and "conventional" as essential characteristics of language seems, so far as I can discover, to have entered modern linguistics through Saussure, we might have expected to find some aid in his writings, but the closest he comes to providing a definition of arbitrariness is "unmotivated" (Saussure 1995: 101), which leaves us in more or less the same place.

<sup>15</sup> Davidson's anomalous monism has aroused not only a great deal of controversy, but also considerable divergences in interpretation. Attempting to examine it with care would require a



The general idea of the following paragraphs will then be to offer a Davidsonian explanation as to why the physical or material aspect of language cannot be used to predict what we mean with a certain word (more specifically, why there can be no strict nomological connections between physical processes and meaning), which should cover an instinctive idea of what arbitrariness means.

Davidson offers his theory of anomalous monism as an attempt to reconcile three apparently true propositions: 1) Mental events (at least some) are causally related to physical events; 2) Causality is determined by strict laws, and 3) Mental events are not subject to strict deterministic laws. The problem with accepting the truth of all three stems from the nature of the interaction between the physical and the mental: if a physical event is caused by a mental event, then it would seem that there are physical events that arise without being reducible to strict causal explanations (which goes against proposition 2), since such an explanation would go against principle 3. (Davidson 1970: 208-209).

Davidson's first move towards a solution is the acceptance of monism, understood as the belief that all events are physical, and mental events are supervenient upon them.<sup>16</sup> In Davidson's case, supervenience means that

---

lengthy investigation of its own, and so I will here limit myself to a general outline of his position, without evaluating his arguments or trying to solve issues of interpretation that will not affect the general purpose of the present essay.

<sup>16</sup> This is a very strong initial assumption and then only argument in its favor I can find in Davidson's works seems to be circular, given it depends on the anomalousness of mental events: if all physical events are strictly determined causally, and there is no strict causation between mental and physical events, then whenever we say a physical event is caused by a mental event there must be a physical description of the mental event that provides the necessary strict causal determination, which means

mental descriptions (which he defines as descriptions that essentially contain a psychological verb that creates a non-extensional context) can be made of at least a certain number of physical events, and that these mental descriptions cannot be restated using a purely physical vocabulary <sup>17</sup> (1970: 210-211; 1973b: 253-254). As an example of how such a supervenience would work, Davidson mentions Tarski's proof that a formal language L cannot contain a predicate stating the truth of its propositions (i.e., cannot contain the predicate "true-in-L") without becoming contradictory; thus, we need a new language (say, L\*) that contains the predicate "true-in-L" in order to characterize what all true propositions in L have in common. The important point here is that this new language does not increase the number of truths in L, but merely offers a redescription of propositions that were already contained in our initial language (1970: 214-215; 1973b: 249-250). In a similar way, mental events are

---

all mental events must be related to a physical one (Davidson 1970: 224). Unless we accept anomalous monism beforehand, we might simply reply that all this argument does is point out the paradox already mentioned, and since proving anomalous monism requires that we previously accept monism, not much seems to be gained with this proof. It seems better, then, to accept monism as a presupposition, and view the whole of anomalous monism as proof that it is a tenable position, even if one that has not been exhaustively demonstrated.

<sup>17</sup> Once again, it might seem this is too strong an idea to be taken for granted; Davidson seems to believe the already mentioned historical failure of reductionism is proof enough, and the only other arguments he offers, related to the different constitutional forces of the physical and the mental, are, or so I believe, inextricably bound with the arguments for anomalous monism itself (*cf.* Davidson: 1973b).

descriptions that depend on predicates that are not contained in the purely physical vocabulary, but do not increase the number of events in the world; in other words, every mental event is a redescription of an event that can be individualized using physical vocabulary, but the information given to us by mental predicates cannot be obtained from the physical predicates that apply to the same event. Thus, still assuming every physical event is strictly determined causally, we can predict the occurrence of a certain event, but this does not imply we can predict its mental description (1970: 215; 1974b: 230).

We must now return to mental predicates: they are, as we said, psychological verbs, such as *believes*, *thinks*, *feels*, *knows*, etc. According to Davidson, attributing any one of these actions to a person implies judging her actions in accordance with what he calls the constitutive idea of rationality, more or less in the same way as describing a physical object forces us to work within the framework of certain constitutive ideas (such as the transitivity of length and weight), and just as we are incapable of even imagining a set of objects the length of which denied the principle of transitivity (i.e., a set of objects *a*, *b*, and *c* such that *a* is longer than *b*, *b* is longer than *c*, but *c* is longer than *a*), we cannot describe a person's behavior using a mental verb unless we are ready to explain that verb as depending on and being coherent with a set of other mental events. In other words, we cannot understand the sentence "X believes *y*", for example, unless we also know that X holds a set *z* of beliefs that are coherent with *y*. And if, later on, we happen to discover X also believes *w*, which is not coherent with the original set *z* we postulated, we must look for a set that is coherent with both *y* and *w*. This, then, is the constitutive force of rationality (1970: 221-222; 1973b: 230-232; 1974b: 236-237).

All of the above becomes relevant for our problem once we realize the constitutive forces of the mental and the physical require that we apply different criteria, and that the constitutive force of the mental, specifically, requires constant changes in the set of beliefs we attribute to a person that have nothing to do with his physical constitution. The two kinds of descriptions we can make of a single event, then, depend on criteria that are unrelated, and one might have to change without our having to modify the other. This, of course, would be impossible if there were such a thing as a strict nomological connection between mental and physical events (1970: 222-223; 1973b: 239). Applying this to the specific case of language, there can be no strict causal law relating what a word means (a mental event) and its sound-pattern (a physical event) without any connection to the language as a whole, or any way of determining what a person means with a linguistic utterance by examining the causal chain of which her utterance is the conclusion. This might in general serve as a rather hazy definition of what arbitrariness is, but let us rephrase it so that it comes closer to our current needs: if we say language is arbitrary, we refer to the fact that there is no strict nomological connection between its meanings and its sound-patterns (i.e., its mental and physical descriptions), and the absence of said nomological connection means we cannot establish the relation between them as necessary, or, given one description, deduce the other. In other words, the only possible proof of the existence of a determinate arbitrary relation is the relation itself: the only way to demonstrate that the word “dog” means dog is by pointing put to a meaningful use of the word “dog” in which a dog is meant. The double use of words related to meaning might seem redundant, but highlights the conclusion of this section, which is that discussion of what a word means

forces us to rely on principles that are neither reducible to physical causality nor follow the same kind of logic.

Now, returning to Davidson's philosophy of language, T-sentences are relations between sound-patterns and meanings such as those mentioned above. But, according to what was proven in section II, they are relations with a law-like behavior<sup>18</sup> that must be verifiable empirically. Given law-like behavior implies counterfactual testing, there is at least the theoretical possibility, even if no such thing actually occurs, that a certain regularity manifest itself in the speaker's behavior. The actual occurrence of this regularity is reinforced by the need for empirical verification: if each sentence were uttered no more than once (assuming, for the moment, that the principle of compositionality does not apply), or if its meaning changed completely from utterance to utterance, the interpreter would have no way to test the validity of the hypothetical meaning she attributed to the sentence upon her first having heard it, and thus she would be unable to justify her hypothetical meaning as an interpretation.<sup>19</sup> The only situation in which an interpreter would appear to be justified in her interpretation of a sentence the speaker utters only once would be that in

---

<sup>18</sup> Within the context of a language, of course.

<sup>19</sup> Davidson never denied the need for this regularity, which he accepted in his reply to Dummett's criticisms (1994: 123), and it continues to show up even when his interest was centered on triangulation, which requires that the interpreter generalize from perceived similarities in the other's linguistic reactions to the external world (*cf.* 1991: 212-213; 1992: 117). It is interesting to notice that Bilgrami, whose project also includes the denial that conventions are philosophically relevant for understanding language, also accepts the need for regularity, although inspired by reasons similar to Davidson's learnability argument (1992: 108-109).

which she had the certainty that the speaker and she found exactly the same circumstances relevant, and that what the interpreter found salient in those circumstances was exactly what the speaker found salient in them, but this would imply the interpreter had access to the speaker's thoughts and, as mentioned above, Davidson believes this can occur only through the intermediation of language, so the possibility turns out to be a *petitio principii*.

From the argument given above in favor of the need for regularity two other conclusions can be reached. First, the interpreter must have knowledge, of whatever kind, of the regularity in the speaker's linguistic behavior. In the absence of said knowledge, the interpreter's passing theory would have nothing towards which to converge, and interpretation would not occur. Second, since, once again, interpretation cannot take place after only one utterance, and perception of the regularity is necessary in order to progress from hypothesis to interpretation, it follows that perception of regularity must be prior to actual interpretation. This second point can be reinforced by appealing to the definition of arbitrary relations proposed previously: since the existence of arbitrary relations cannot be a necessary consequence of anything else and must be demonstrated by direct attestation of the relation, the fact that an utterance is a meaningful sentence (that is to say, that there is a law-like T-sentence that determines a relation between truth conditions and sound-pattern) and not random noise can only be proven by direct attestation of its law-like characteristics, and this implies the existence of regularity.

Let us sum up, rephrasing it slightly, what we have concluded so far. If linguistic communication is to succeed, the speaker's and interpreter's passing theories must converge. Unless we accept that the interpreter has previous knowledge of the speaker's language that is

necessary for communication, convergence of passing theories can occur only through radical interpretation. Radical interpretation, in turn, requires that the interpreter know, in some sense of knowing, that the speaker's linguistic behavior is law-like. Knowledge of the existence of said law-like behavior must be prior to actual interpretation, its existence cannot be proven except by direct attestation of the behavior as such, and its specific characteristics cannot be determined by any non-linguistic fact. Lastly, given the relation established by the behavior is arbitrary, a different set of connections between sound-patterns and truth conditions would work equally well. I believe that the similarity between these conditions for successful linguistic communication and the six characteristics of conventions proposed by Lewis<sup>20</sup> is noticeable enough to consider our conditions a kind of convention (the most notorious differences will be discussed in section IV). Thus, if my analysis so far is correct, we could reverse (or, rather, return to its usual direction) Davidson's conclusion (*cf.* Davidson 1982: 280): linguistic communication is not a condition for having conventions. Rather, conventions are a condition for linguistic communication.

---

<sup>20</sup> They can be summarized as follows: a convention is a regularity R such that: 1) everyone conforms to R; 2) everyone believes that the others conform to R; 3) this belief gives everyone a good reason to conform to R; 4) there is a general preference for general conformity to R rather than slightly less than general conformity; 5) R is not the only possible regularity meeting the last two conditions; and 6) the various facts listed in 1-5 are matters of general knowledge (Lewis 1983: 165-6).

## IV

I can think of three counterarguments that can be raised, from within Davidson's theories, against the conclusion in section III.<sup>21</sup> The first has to do with conventions not being

---

<sup>21</sup> An anonymous reviewer for *Manuscrito* has made me notice another possible objection, although one more related to the exegesis of Davidson than to the argument itself (hence its being in a footnote); "A Nice Derangement of Epitaths" is not meant to be read in conjunction with Davidson's earlier work, but with his later philosophy and its denial of clear-cut boundaries between our knowledge of language and our knowledge of the world. An example of this would be part of the conclusion from "On the Very Idea of a Conceptual Scheme" where Davidson defends the thesis that interpreting a speaker requires attributing to her true beliefs (according to our own knowledge) that are relevant in the context of utterance as the meaning of those sentences she holds true in order to begin the process of understanding her language, and given we attribute beliefs on the basis of linguistic skills, there would be no way of recognizing a speaker without assuming she holds a set of beliefs similar to our own (1974c: 196-7: cf. 1974a: 152-154: 1975: 168-70). This, of course, means that meanings exist only as part of our general knowledge of the world, as part of the interconnected set of our beliefs (or, to use Pietroski's words, communication depends on General Intelligence) (Pietroski 1994: 104-105). Without entering into a detailed discussion of the arguments behind these ideas, there are two things I believe should be pointed out: the first is that the reliance on the Principle of Charity seems to point the way towards the mechanics of radical interpretation (which show up quite often throughout the three essays already mentioned), which would mean that, as far as I can see, there is no clear break in Davidson's thought. The second is that the lack of a clear distinction between meaning and knowledge of the world does not mean language cannot depend on conventions. In fact, most any convention requires knowledge of the world in order to be applied correctly (there is no such a thing as a convention that



truly prior to linguistic communication. The second stems from how reduced, both in time and extension, conventions would turn out to be. The third is related to the philosophical relevance of such conventions. I will try to answer them one by one, although my answers will be meant to show that the issue has not been conclusively settled, even within Davidson's theories, rather than to close the discussion.

1) The need for conventions prior to successful interpretation has already been discussed; nevertheless, it could be argued that such conventions arise only after the actual encounter between interpreter and speaker has begun, and thus cannot be learned beforehand (*cf.* Davidson 1982: 277-278). The conflict between both positions can be better grasped, I believe, by asking when linguistic communication begins: if it is simultaneous with the encounter, then the counterargument is correct, and convention would be posterior to linguistic communication. On the other hand, if linguistic communication is only said to begin once successful interpretation occurs, then the analysis offered in section III still holds, given the interpreter has to acquire knowledge of a certain number of T-sentences prior to an interpretation that she can consider reliable, and thus may be said to have learned these conventions before linguistic communication took place. This might initially seem a verbal quibble, but I believe there is a good philosophical reason for believing that linguistic communication begins with interpretation. Intuitively, at least, if given two sets, the first consisting of entities (understood in a very wide sense) such as tables, chairs, solar eclipses, and magpies, and the

---

specifies beyond all possible doubt every single aspect of its execution or the moment when it should be applied) (*cf.* Millikan 1998: 14-16).

second consisting of entities such as bets, flirtations, weddings, and baptisms, we would classify linguistic communication in the second set. Now, part of what characterizes the entities in the second set is that they can fail, while those in the first cannot. The question is, then, is a failed wedding a kind of wedding? Once again, intuitively, the answer would be that it is not. So would failed linguistic communication be a kind of linguistic communication? We might be tempted to say that, if the wrong meaning is transmitted, it might be, but if no meaning at all is transmitted, then it is not.<sup>22</sup> Since, as has been stated repeatedly, no meaning is transmitted until regularity sets in, we might say linguistic communication can only begin after regularity is perceived, which would be the same as saying linguistic communication begins after a convention is established. These intuitions, of course, require careful examination before they can be considered valid, but at least they seem to indicate the question has not been settled.<sup>23</sup>

---

<sup>22</sup> This answer is, quite evidently, heavily indebted to Austin's work on infelicities (1962: 14-24).

<sup>23</sup> Armstrong has offered a reding of Lewis that replies provides us with an argument similar to my own; conventions do not require the existence of past occurrences, but the commitment to follow a similar behavior and the expectancy of it should a similar situation arise (Armstrong 2016; 99-100). Although I agree with, for example, Millikan, who argues that conventions need not be repeated every time, or even most times, but simply enough times for their survival to be guaranteed (that is to say, enough times for the speaker to find it worthwhile to try it out, hoping thus to increase the chances for her desired outcome) (Millikan 1998: 14-16; 2003: 41-43), so that a weak reading of regularity suffices for conventionality, as indicated in section III, some actual regularity is needed in order to interpret something as a linguistic convention (otherwise, there would be no way to verify the law-

2) This second objection can actually be subdivided into two parts.<sup>25</sup> First, the kind of convention my conclusion proposes might end up applying to no more than two people and for no longer than a single conversational encounter; does such a small regularity deserve the rather large name of “convention”? In fact, we relied on a similar argument to discard passing theories as languages in section I. Can we justify the change of attitude without being inconsistent? I think the latter question should be answered affirmatively: back then, we lacked a considerable amount of information concerning what passing theories relied on, and thus had no very good reason for believing convergences had any serious connection to conventions. But, if the conclusion of section III is correct, we know now that they share a considerable number of similarities, and we are thus entitled to reconsider the validity of our initial argument. The point, then, boils down to the first question. Does the number of people sharing it and how long it lasts determine whether something is a convention? The question might be better phrased as follows: is there an essential difference between a convention shared for a short period of time by only two people and one with a greater temporal and social extension? Let us consider the temporal aspect first. If there were an essential difference, we would have to say the convention’s existence is causally determined by a temporal element, which would mean, as

---

like T sentences), so I cannot agree with Armstrong, at least when dealing with the scenario Davidson is considering, where interpreter and speaker have no previous channel for linguistic communication. If this is not the case, one-time coordination might occur, but the conventions would have been discussed and determined prior to the first encounter.

<sup>25</sup> Armstrong discusses a similar issue as part of his defense of Lewisian conventionality (Armstrong 2016: 99-100).

mentioned in section III, the convention would no longer be arbitrary, and this is absurd. A similar argument applies for its social extension. From a positive point of view, the argument would state that what is essential to a convention is the possibility of externally unmotivated regularity and of being shared, not the actual fact of reappearing a determinate number of times or being shared by a certain number of people.

The last point brings us to the second part of this counterargument: if Davidson's account of communication nowhere requires that speaker and interpreter use the same language, but only that they understand each other's language, can we still say these supposed conventions are shared (*cf.* Davidson 1992: 14)? The answer to this takes us deeper into the notion of regularity: is there regularity only when two people, for example, behave in the exact same way? Or would we say there is a regularity when, every time one does x, the other does y? I see no reason to deny the latter is as much a regularity as the former, and as long as it complies with other requirements (such as arbitrariness), it would be a convention. Consider, for example, courtly etiquette: every time the king, let us say, enters the room, his steward bows to the east. Both of them have a different role, but the action as a whole is both a regularity and a convention. In a similar vein, we would say that whenever the speaker utters s, her interpreter understands p: it is the action as a whole, not what each participant does, that must be understood as a regularity.<sup>26</sup>

Before proceeding to the last counterargument, a clarification is in order. Does my reply mean that passing theories are languages? If passing theories are meant to include interpretation of malapropisms and other mistakes

---

<sup>26</sup> Millikan follows a similar line, though in a slightly different context (Millikan 1998: 4-6; 2005a: 58-60).

that occur but once, then the answer is no, given there is no conventional regularity. And if we admit, as discussed in section I, that the conventionality of language does not mean it is completely conventional, this does not represent a problem.

3) The last counterargument is apparently similar to the second: if these supposed conventions apply at such small scale, and given each individual has a different vocabulary and set of syntactical habits, our notion of language would be reduced to either an idiolect, or a haphazard collection of idiolects, both of which turn out to be philosophically uninteresting (*cf.* Davidson 1994: 111, 115). These comments were originally meant as replies to Dummett's critical remark concerning the mistaken priority given to idiolects in Davidson's philosophy (Dummett 1986: 462); the debate between Dummett and Davidson is, in itself, a philosophical problem, and thus, without wanting to commit myself to Dummett's position, I would like to endorse what seems to be the general idea behind his remarks. It is true that the conventions mentioned in this essay might appear as idiolects, but only assuming we define an idiolect as what only one person actively speaks; if the reply to (2) is correct, then even an idiolect must be a regularity shared by at least two people, and nothing denies the possibility of its being shared by many more. We might restate this by saying Davidson's view emphasizes the actual outcome, while my own view (and, perhaps, Dummett's) favors what is possible. Adopting the latter point of view would mean considering idiolects languages that happen to be shared by only two people. The reason to prefer this position has to do precisely with the philosophical significance of what is potentially contained in language. As an example, let us consider writing: even if, assuming Davidson is correct, communication is possible because interpreter and speaker share a context against

which they can determine the truth conditions of their utterances, once a linguistic convention is established, language can not only overcome its own initial conditions (by separating the author's from the reader's context), but even be used to create new ones, within which its initial conventions can be altered.<sup>27</sup> I cannot offer a precise criterion for philosophical relevance or significance, but am nevertheless sure this is a subject well worth any philosopher's time.

If my general analysis of Davidson's theories is correct, we can now conclude that, unless an account quite different to radical interpretation can be made to explain the convergence of passing theories, conventions are necessary in order to account for the very possibility of linguistic communication. As far as I can see, accepting them does not prove detrimental to the rest of Davidson's account, and examining the role they are meant to play, as was done indirectly in this section, might allow us to apply many of Davidson's insights to subjects that, initially, seem external to his main interests. At the very least, I hope to have demonstrated how many issues are still open to debate even within Davidson's theories, and hinted at what might be gained by looking for answers.

## REFERENCES

ARMSTRONG, J. Coordination, Triangulation, and Language Use. *Inquiry* 59(1), 80-112, 2015.

\_\_\_\_\_. The problem of lexical innovation. *Linguistics and Philosophy* 39(2), 87-118, 2016.

---

<sup>27</sup> Judging by essays such as *Locating Literary Language*, Davidson was not only aware, but also acutely interested in such problems.

- AUSTIN, J.L. *How to Do Things with Words*. Great Britain: Oxford University Press, 1964.
- BILGRAMI, A. (1992) *Belief and Meaning: The Unity and Locality of Mental Content*. Cambridge, MA: Blackwell Publisher, 1992.
- CAIN, M.J. (2013) Conventions and Their Role in Language. *Philosophia* 41, 137-158, 2013.
- DAVIDSON, D. (1964) Theories of Meaning and Learnable Languages. In: Bar-Hillel, Y. (ed.). *Proceedings of the International Congress for Logic, Methodology, and Philosophy of Science*. Amsterdam: North-Holland. Reprinted in: Davidson (1984) (pages 3-15).
- \_\_\_\_\_ (1967) Truth and Meaning. *Synthese* 17. Reprinted in: Davidson (1984) (pages 17-36).
- \_\_\_\_\_ (1970) Mental Events. In: Foster, L. & Swanson, J.W. (eds.), *Experience and Theory*. London: Duckworth. Reprinted in: Davidson: 2001b (pages 207-225).
- \_\_\_\_\_ (1973a) Radical Interpretation. *Dialectica* 27. Reprinted in: Davidson (1984) (pages 125-139).
- \_\_\_\_\_ (1973b) The Material Mind. *Studies in Logic and the Foundations of Mathematics* 74. Reprinted in: Davidson (2001b) (pages 245-259)
- \_\_\_\_\_ (1974a) Belief and the Basis of Meaning. *Synthese* 27. Reprinted in: Davidson (1984) (pages 141-154).
- \_\_\_\_\_ (1974b) Psychology as Philosophy. In: Brown, S. (ed.). *Philosophy of psychology*. New York: Harper & Row. Reprinted in: Davidson (2001b) (pages 229-239).

- \_\_\_\_\_ (1974c) On the Very Idea of a Conceptual Scheme, *Proceedings and Addresses of the American Philosophical Association* 47, 5-20. Reprinted in: Davidson (1984) (pages 183-198).
- \_\_\_\_\_ (1975) Thought and Talk. In: Guttenplan, S. (ed.) *Mind and Language*. Oxford: Oxford University Press. Reprinted in: Davidson (1984) (pages 155-170).
- \_\_\_\_\_ (1976) Reply to Foster. In: Evans, G. & McDowell, J.H. (eds.) *Truth and Meaning: Essays in semantics*. Oxford: Clarendon Press. I Reprinted in: Davidson (1984) (pages 171-179).
- \_\_\_\_\_ (1982) Communication and Convention. *Synthese* 59. Reprinted in: Davidson (1984) (pages 265-280).
- \_\_\_\_\_ (1984) *Inquiries into Truth and Interpretation*. New York: Oxford University Press.
- \_\_\_\_\_ (1986) A Nice Derangement of Epitaphs. In: LePore (1986). Reprinted in: Davidson (2005) (pages 89-107).
- \_\_\_\_\_ (1991) Three Varieties of Knowledge. In: Phillips Griffiths, A. (ed.), *A.J. Ayer Memorial Essays: Royal Institute of Philosophy Supplement*, 30, Cambridge: Cambridge University Press. Reprinted in: Davidson (2001a) (pages 205-220).
- \_\_\_\_\_ (1992) The Second Person. *Midwest Studies in Philosophy* XVII. Reprinted in: Davidson (2001a) (pages 107-121).
- \_\_\_\_\_ (1993) Locating Literary Language. In: Dasencrock, R.W. (ed.) *Literary Theory After Davidson*. University Park: Pennsylvania State University Press. Reprinted in: Davidson (2005) (pages 167-181).
- \_\_\_\_\_ (1994) The Social Aspect of Language. In: McGuinness, B. & Oliveri, G. (eds.). *The Philosophy of*



- Michael Dummett. Dordrecht: Kluwer academic Publishers.  
Reprinted in: Davidson (2005) (pages 109-125).
- \_\_\_\_\_ (1999) Reply to James Higginbotham. In: Hahn, L.E. (ed.), *The Philosophy of Donald Davidson* (pages 687-689). Illinois: Open Court.
- \_\_\_\_\_ (2001a) *Subjective, Intersubjective, Objective*. New York: Oxford University Press.
- \_\_\_\_\_ (2001b) *Essays on Actions and Events*. Oxford: Oxford University Press.
- \_\_\_\_\_ (2005) *Truth, Language, and History*. New York: Oxford University Press.
- DUMMETT, M. A Nice Derangement of Epitaphs: Comments on Davidson and Hacking. In: LePore (1986) (pages 459-476), 1986.
- GREEN, K. Davidson's Derangement: Of the Conceptual Priority of Language. *Dialectica* 55(3), 329-258, 2001.
- GOLDBERG, S.C. Radical Interpretation, Understanding, and the Testimonial Transmission of Knowledge. *Synthese* 138, 387-416, 2004.
- HACKING, I. (1986) The Parody of Conversation. In: LePore (1986) (pages 447-458).
- LEPORE, E. (ed.) (1986) *Truth and Interpretation. Perspectives on The Philosophy of Donald Davidson*. New York: Basil Blackwell.
- \_\_\_\_\_ & Ludwig, K. *Donald Davidson. Meaning, Truth, Language, and Reality*. New York: Oxford University Press, 2005.
- LEWIS, D. (1983) Language and Languages. In: Gunderson, K. (ed.). *Minnesota Studies in the Philosophy of Science*.

*University of Minnesota Press*. Minneapolis: University of Minnesota Press. Reprinted in: Lewis, D., *Philosophical Papers. Volume I* (pages 163-188). New York: Oxford University Press.

LUDLOW, P. *Living Words. Meaning Underdetermination and the Dynamic Lexicon*. Oxford: Oxford University Press, 2014.

MILLIKAN, R.G. (1998) Language Conventions Made Simple. *Journal of Philosophy* 95(4), 161-180.  
Reprinted in: Millikan (2005b) (pages 1-23).

\_\_\_\_\_ (2003) In Defense of Public Language. In: Antony, L. N. & Hornstein, L. (eds.) *Chomsky and his Critics*. Oxford: Wiley-Blackwell (pages 215-237). Reprinted in Millikan (2005b) (pages 24-52).

\_\_\_\_\_ (2005a) On Meaning, Meaning, and Meaning. In: Millikan (2005b) (pages 53-76).

\_\_\_\_\_ *Language: A Biological Model*. Oxford: Oxford University Press. 2005b.

PIETROSKI, P. A Defense of Derangement. *Canadian Journal of Philosophy* 24(1), 95-117, 1994.

REIMER, M. What Malapropisms Mean: A Reply to Donald Davidson. *Erkenntnis* 60, 317-334, 2004.

DE SAUSSURE, F. *Cours de linguistique générale*. Paris: Éditions Payot & Rivages, 1995.

STANTON, R.J. A Deranged Argument against Public Languages. *Inquiry* 59(1), 6-32, 2016.

