6

Betraying Davidson: A Quest for the Incommensurable[1]

Marcello Di Bello

6.1 Introduction

Talks of incommensurable conceptual schemes typically allude to the following picture: There is *reality*[2] that is inaccessible *per se*, on the one hand, and there is us, or different groups of people, accessing reality by means of *conceptual schemes*, on the other. As long as schemes radically differ from each other, they will stand to one another in a relation of incommensurability. But what does it mean for two schemes to be radically different and thus be in a relation of incommensurability? And is such a relation at all intelligible to us?

Davidson (1974) has formulated an argument to the effect that incommensurability is unintelligible.[3] More precisely, his argument first provides a definition of incommensurability, and then shows that incommensurability thus defined is unintelligible. I will argue, *contra* Davidson, that his definition of incommensurability does not lend any support to the unintelligibility claim. To be sure, my attack against Davidson will be in three forms. For I will give three interpretations of his argument against incommensurability, and then provide three counter-arguments, which will have progressively weaker strength. The

[1]I would like to thank in particular Martin Stokhof, Michal Lukasiewicz, Sven Lauer, and Ansten Mørch Klev.

[2]Or any other controversial term one may prefer to use for what is “out there.”

[3]See Davidson (1974).

Language, Knowledge, and Metaphysics.
Massimiliano Carrara and Vittorio Morato (eds.)
College Pubblication, London, 2008.

first refutes the argument that emerges from the first interpretation; the second can do so with respect to the argument that emerges from the second interpretation on the basis of additional assumptions; finally, the third provides only a conjecture that the argument that emerges from the third interpretation fails.

The plan for the paper is as follows. Section 6.2 points out distinctions and inherent difficulties that are peculiar to the incommensurability debate. Section 6.3 reconstructs Davidson's definition of incommensurability. Second 6.4 contains Davidson's argument against the intelligibility of incommensurability, along with its three interpretations. Section 6.5 contains my three counter-arguments. Finally, section 6.6 concludes and points to open problems.

6.2 Preliminaries

Anyone who joins the debate about incommensurability should bear in mind certain distinctions and inherent difficulties. I would like to point out two distinctions and one difficulty, without intending to be exhaustive. The first distinction is the one between total and partial incommensurability. Intuitively, partial incommensurability holds only between restricted regions of two supposed incommensurable schemes, while total incommensurability holds between the schemes as a whole.

The second distinction is among different forms of incommensurability—for instance, semantic, methodological, and ontological incommensurability.[4] In this paper, I will be concerned with *semantic* incommensurability only. Roughly speaking, semantic incommensurability arises between languages *qua* carriers of meaning and indicates a radical meaning-difference between languages. What such a radical meaning-difference between languages is will be made more precise in the next section.[5]

Evidently, the notion of meaning is to play a crucial role for defining semantic incommensurability. It follows that the well-known difficul-

[4] See Sankey (1993).

[5] It is useful to contrast semantic incommensurability with ontological and methodological incommensurability. Incommensurable conceptual schemes can be taken to be languages carrying meaning, and this perspective suits the notion of semantic incommensurability. Conceptual schemes, however, can be taken to express the ontology adopted by a community of speakers. Ontology here means what is taken to exist and what not, how classifications of entities are carried out, or how the plethora of being, so to speak, is divided up. This perspective suites the notion of ontological incommensurability. Further, conceptual schemes can be taken to be scientific theories which are associated with standards and criteria concerning theory appraisal, *e.g.*, concerning what counts as evidence in favor or against the theory. This other perspective pairs with methodological incommensurability.

ties in singling out identity criteria for meaning are inherited by any attempt to define semantic incommensurability. For the sake of clarity, in the rest of the paper I will attempt to be very explicit about which assumptions about meaning I have endorsed in laying down any criteria for semantic incommensurability.

6.3 A Davidsonian definition of incommensurability

As anticipated, Davidson's argument against incommensurability is in two parts. First, a definition of incommensurability is given, and secondly incommensurability thus defined is shown to be unintelligible. This section discusses the moves made in the definitional part of the argument in which Davidson's final definition of semantic incommensurability reads: (INC-D) *Two schemes are semantically incommensurable iff (i) they fail to be intertranslatable and yet (ii) they are both true.*[6]

In what follows I will explain the significance of points (i) and (ii) of definition (INC-D). It is not my concern to defend definition (INC-D) in full, but only to elucidate it. In doing so, however, it will become apparent that (INC-D) should be slightly modified.

6.3.1 Failure of intertranslatability between languages

The definitional moves suggested by Davidson to motivate point (i) are as follows:

(M1) A conceptual scheme can be identified with a language.[7]
(M2) Untranslatability between languages is evidence for incommensurability.[8]

Move (M1) makes it precise that a linguistic-based incommensurability is at stake. As said before, the focus of this paper is semantic incommensurability, and it is natural to think of semantic incommensurability as a relation between languages *qua* carriers of meaning. If one were to deny that semantic incommensurability involves languages, one will have to deny that meaning is related to, or resides in, languages.

I did not yet say what a language should be taken to be. A language is customarily defined as the set of all well-formed sentences which

[6] "And the criterion of a conceptual scheme different from our own now becomes: largely true but not translatable." (Davidson, 1978, p. 194). Note that here and in the following all quotations and page numbers are in accordance with Davidson (1978), which is a collection of of all the Davidson's papers I will refer to in this paper.

[7] "We may identify conceptual schemes with languages." See Davidson (1978, p. 185).

[8] "Can we then say that two people have different conceptual schemes if they speak languages that fail of intertranslatability?" See Davidson (1978, p. 185).

can be constructed out of a given alphabet and a grammar. However, since a language is here intended to be useful for defining semantic incommensurability, in this context it cannot consist of syntactic entities only, for these must be paired with a semantics. Without any semantics in place the notion of semantic incommensurability between languages would be empty. Yet syntax and semantics are not enough to define a language, for a language is *used* by a community of speakers to communicate. Thus, here I take a language to be the set of well-formed sentences each endowed with meaning and used by a community of speakers to communicate.

Move (M2) is a natural follow-up to move (M1): If the *relata* in the relation of incommensurability are languages rather than schemes, the criterion for them to be radically different from each other, and thus incommensurable, is the failure of intertranslatability, or the non-intertranslatability, between the languages. Given the switch from schemes to languages, the failure of intertranslatability is the most obvious candidate that can function as evidence for radical meaning-difference between languages, and thus for semantic incommensurability.

To make the notion of non-intertranslatability precise, a few definitions are in order. The non-intertranslatability of two languages L and L' comes in two forms, partial and total. Languages L and L' are *partially* non-intertranslatable iff, for *some* sentences in L, a meaning-equivalent sentence in L'cannot be found, and viceversa. Languages L and L' are *totally* non-intertranslatable iff, for *all* sentences in L, no meaning-equivalent sentence in L' can be found, and viceversa. Depending on whether the failure of translatability is partial or total, incommensurability will be partial or total.[9]

6.3.2 Introducing the notion of truth

The definition of incommensurability that would result from moves (M1) and (M2) alone—*i.e.*, two languages are incommensurable iff they are non-intertranslatable—is insufficient. To see this, let us forget for a

[9] In the definition of (non-)intertranslatability the notion of meaning-equivalence was invoked, and yet it was left undefined. This is one of the point in the discussion about semantic incommensurability in which an appeal to different theories of meaning, and thus to different identity or equivalence criteria for meaning, can make a difference. Davidson does not make explicit which identity criteria for meaning he adopts. However, given that Davidson's account of meaning can be roughly seen as truth-conditional, for him the criterion of meaning-equivalence between sentences is likely to be truth-conditional equivalence. So, in my critique of his argument I will follow him with respect to this assumption.

moment the switch from schemes to languages, and let us concentrate on some features of conceptual schemes. Davidson makes two points about the role of conceptual schemes. First, schemes are adopted by a community of agents or speakers to categorize reality; they are about an unschemetized content which lies beyond the schemes themselves. Secondly, schemes are adopted by a community because they satisfy a certain normative requirement, *i.e.*, they fit the available evidence. But—Davidson concludes—when a scheme fits the available evidence, this means that a scheme is *true*. Now, if we grant that schemes are to be true to count as schemes, it follows that if two schemes stands in a relation of incommensurability both schemes are true.[10] So, the final definition of incommensurability will boil down to (INC-D).

However, as soon as the notion of truth is applied to languages, rather than schemes, a problem arises. In the wording of definition (INC-D), Davidson has maintained that non-intertranslatablity is evidence for incommensurability and that schemes should be true, but he has forgotten move (M1), namely that schemes can be replaced by languages. But note what happens if the contribution of (M1) enters definition (INC-D). The result would be as follows: *Two languages are semantically incommensurable iff (i) they fail to be intertranslatable and yet (ii) they are both true.* Clearly, this definition is absurd because languages cannot be true.[11]

Allegedly, Davidson had implicitly assumed that some subsets of the chosen languages, not the entire languages, are taken to be true. The main, and non trivial, problem is *which* subsets of the given languages are to be taken to be true. My suggestion is along these lines. Suppose we fix a state of the world at a certain time and place. Then, speakers

[10] A further argument for invoking the notion of truth is this. Let us reason for a moment in terms of scientific theories, which can be thought of as highly formalized conceptual schemes in which scientific evidence fits. Now, the incommensurability of two scientific theories presupposes that there is no available criteria to prefer one over the other. For the sake of argument, suppose the requirement of both theories/schemes being true is dropped from the definition of incommensurability. Then, we would have three cases: both theories are false; the truth-value of the theories is unknown; and one theory is true and the other is false. With respect to the first two cases, none of the theories would be even endorsed by any scientific community, because of its falsity or because of the patent lack of supporting evidence. In the third case, there would be a criteria to discriminate between the two theories, namely dismissing the false theory and retaining the true theory. In either case, incommensurability would vanish. The fact that both theories are true, thus, preserves the indecision which is required for incommensurability to obtain.

[11] Some commentators, *e.g.* Hacker (1996), go on saying that if Davidson holds that languages are true, then he has to hold that contradictions are true, which is absurd. However correct, I believe that this is a critique which is directed against the letter, but not the spirit of Davidson's argument.

of the two supposed incommensurable languages will hold true certain subsets of their languages. By a Davidsonian principle of charity, the two subsets that are *taken to be true* by the two communities of speakers are regarded as simply *true*. In light of this observation, the amended version of definition (INC-D) I shall propose reads: (INC-D*) *Two languages are semantically incommensurable iff (i) they fail to be intertranslatable and yet (ii*) subsets of them are true (those subsets that are taken to be true by the communities of speakers associated with the two languages).*

6.4 Davidson against incommensurability

Having spelled out what incommensurability is for Davidson, it is time to reconstruct his argument against it. I have said that the objective of Davidson's argument is the claim that incommensurability is unintelligible, or that such a notion cannot be made sense of. What does that mean? It is instructive to distinguish three claims:

(1) It is not possibile that incommensurability holds.
(2) Members of any community of speakers cannot coherently claim that their language is incommensurability to another.
(3) Members of our community of speakers cannot coherently claim that our language is incommensurable to another.

Preliminarily, incommensurability holds iff there exists a pair of languages (L_1, L_2) such that L_1 and L_2 stand in the relation defined by (INC-D*). So (1) is a modal claim about the *impossibility* of incommensurability, while (2) and (3) are claims about the *speakers' inability* of asserting the possibility of incommensurability. But which of these claims is the objective of Davidson's argument? I have no definitive answer to this question, so in what follows I will give three readings of the argument. One reading takes Davidsons to claim (1), the other takes him to claim (2), and yet another takes him to claim (3).

I shall begin by reconstructing the common ground of Davidson's argument, and then launch the three different readings. Incommensurability entails that two languages (or rather, some properly chosen subsets of them) are both true and yet untranslatable, by clauses (i) and (ii*) of (INC-D*). But what is truth here? Davidson makes use of Tarski's theory of truth, as he believes it to be the most natural account of what we understand by truth.[12] Tarski's theory of truth associates with each sentence φ of a language L a bi-conditional (T) of the form:

[12] Some authors have objected that Tarski's theory of truth is not as innocuous as Davidson would like it to be (see Hacker (1996)). While I believe this is certainly the case, I will not overload my critique with this observation.

(T) φ is true iff $\overline{\varphi}$,

where 'φ' is replaced by a name for φ and '$\overline{\varphi}$' by a translation of φ into the meta-language, such that $\overline{\varphi}$ is a translation of φ.[13]

Given this characterization of what it takes for a sentence to be true, Davidson's argument in his own words runs as follows:

> And the criterion of a conceptual scheme different from our own now becomes: largely true but not translatable. The question whether this is a useful criterion is just the question how well we understand the notion of truth, as applied to language, independent of the notion of translation. The answer is, I think, that we do not understand it independently at all. ... Since convention (T) embodies our best intuition as to how the concept of truth is used, there does not seem to be much hope for a test that a conceptual scheme is radically different from ours if that test depends on the assumption that we can divorce the notion of truth from that of translation (Davidson, 1978, p. 194).

In nutshell the argument says that, while incommensurability requires truth and translatability to *be* independent, as a matter of fact truth and translatability *are not* independent as convention (T) shows. More carefully, the *understanding* of the notion of truth is not independent of the understanding of the notion of translatability, and yet incommensurability would require the two notions to be independently understood. The argument is very brief and needs to be unpacked. In fact, depending on where the emphasis is placed—*i.e.*, on incommensurability as such, or on speakers' inability of understanding incommensurability, or on our inability of understanding incommensurability—three readings of the argument can be given, yielding claim (1), (2) and (3) respectively. I shall consider each reading in turn.

Under the *first* reading, the argument would run as follows. Recall that for languages L and L' to be incommensurable the notion of truth should be invoked. So according to the bi-conditional or convention (T), when L and L' are both true (or better, subsets of them are true), their sentences are both translatable into a common meta-language. Strictly speaking, in Tarski's theory, given two languages, there is no mention of a *common* meta-language in which truth-conditions are spelled out. However, in the case at hand, one wants to say that two languages are *both* true, so a *common* meta-language should be used, otherwise the two languages will be true under different standards. Thus, the two supposed incommensurable languages will have to be both translatable into the same meta-language. But since both languages are translat-

[13]To be precise, in the definition of the T-condition, one has to add that the syntax and semantics of the meta-language should be previously defined.

able into the same meta-language, this suggests that there must be a way to translate sentences of one language into sentences of the other language, and viceversa, by means of the shared meta-language which would function as a common yardstick of meaning comparison that is suited to carry out the translation. In other words, the two supposed incommensurable languages must be intertranslatable. So the claim of languages/schemes being true and yet untranslatable cannot possibly hold. This shows that claim (1) holds.

The first reading implicitly assumes that the claim that languages L and L' are incommensurable would be formulated by a third-language perspective, *i.e.*, the meta-language into which both L and L' can be translated and from which both L and L' can be compared and inspected, so to speak. But the assumption of a third-language perspective is debatable. Some may hold that any incommensurability claim is formulated, or takes place, within one of the two languages which are claimed to be incommensurable. So, the *second* reading of the argument makes it explicit that:

(C) The incommensurability claim takes place within either L or L' (*i.e.*, no third language perspective is allowed).

This is reasonable in the case in which the speakers of a given language want to claim that an alien language is incommensurable to theirs.

Assumption (C) has an interesting consequence, *i.e.*, the language in which the incommensurability claim is formulated will also play the role of the meta-language. This can be easily seen. If the incommensurability claim about L and L' is localized in L, it is within L that the statement that the sentences of L and L' are true should be expressed (recall that incommensurability requires truth). So, it is by using the linguistic resources of L that convention (T) should be expressed. But convention (T) is written in the meta-language, so L is to play the role of both the object-language and the meta-language. Now, from the first reading we know that both languages L and L' must be translatable into the common metalanguage. But the meta-language in this case is L itself. Thus, L' would be translatable into L. As a result, L and L' cannot fall under (INC-D*) because point (i) is not satisfied, hence we cannot coherently claim that L and L' are incommensurable. This shows that claim (2) holds.[14]

The second reading assumes that the claim of incommensurability is made from the standpoint of *any* language. This is not incorrect, but

[14]Note that claim (2) is not that L and L' are not incommensurable, rather that whenever an incommensurability claim is being made in accordance with assumption (C), we find ourselves in a (performative) contradiction.

a situation that is more closely related to us is the one in which the claim of incommensurability is formulated within our own language, *i.e.*, English. So, the *third* reading of the argument can be yielded by replacing condition (C) with the following:

(C*) The incommensurability claim takes place within English.

The argument goes very similarly to the one just given for the second reading: It is enough to replace 'L' with 'English.' Very succinctly, if the claim of incommensurability is made within English, the language L' which is claimed to be incommensurable to English will be translatable into English because English will play the role of the meta-language. Hence, claim (3) holds.

6.5 Betraying Davidson

The task is now to provide a counter-argument to the effect that the validity of the arguments leading to claim (1), (2), and (3) is undermined.[15] Three readings of Davidson's argument have been given, and thus different counter-arguments are needed, depending on the preferred reading. To refute the argument coming out of the first reading, it is enough to be able to construct an instance of (INC-D*) without generating any contradiction. Instead, to refute the argument coming out of the second reading, it is necessary to construct an instance of (INC-D*) such that, in addition, (C) is satisfied. In what follows, I will construct an instance of (INC-D*), which suffices to refute the first reading of Davidson's argument and thus claim (1). My attack against claim (2) and the second reading will be less straightforward. I will show that, under additional assumptions which Davidson should be prone to accept, claim (2) fails.

A separate consideration should be given to the argument for (3). My strategy will be to point out that, although (3) is a special case of (2), considering English instead of some other language does not make any substantial difference. Thus, if (2) fails, there is no evidence for denying that (3) would fail as well.

[15] There is a peripheral worry to be addressed. While Davidson's argument adopts (INC-D) as a definition, the counter-argument I shall give adopts (INC-D*) rather than (INC-D) because, as previously shown, (INC-D) needs to be slightly amended on pain of being incorrect. So a question suggests itself: is the change from (INC-D) to (INC-D*) crucial for the success of my attack against Davidson? This legitimate worry is unavoidable and cannot be fully dismissed. Yet notice that definition (INC-D) as it stands is incorrect (languages cannot be true). Hence, the adjustment from (INC-D) to (INC-D*) was an attempt to read Davidson in the most charitable way. Indeed, the doubt lingers whether a different way to amend (INC-D) would render my counter-argument ineffective, but I shall put this doubt aside.

6.5.1 Against the first reading argument

Consider two different communities of speakers endowed with two different languages.[16] Both communities are concerned with the same phenomenon (or piece of reality, uninterpreted content, etc.), namely the phenomenon of some objects being next to a given object which both communities can clearly identify and refer to by means of a proper name. Obviously the two communities will use different names, but for the present purpose it suffices to refer to such an object by the constant 'c.' The two communities also share the predicate 'next-to' (or some translation of it in the respective languages) whose meaning is the intended one.

However, the communities are not alike in that they have developed different strategies to talk about objects being next to c. One community is only able to say whether there are, or there aren't, objects next to c. The other community, instead, is more precise and can express the exact number of objects next to c, ranging from 1 to any finite number. Despite its precision, the latter community is unable to express the fact that there are 0 objects next to c, probably because such a phenomenon has never been experienced.[17]

My claim is that the languages of the two communities are incommensurable in the sense of (INC-D*). If this can be argued, an instance of incommensurability (INC-D*) can be given as a result, whereby showing that claim (1) is false. Now, recall that definition (INC-D*) is composed of two points. Point (ii*) requires certain subsets of the two languages to be true. The subsets are chosen depending on which sentences the communities of speakers hold true. To satisfy point (ii*), we may suppose that one community holds true the sentence 'there are some objects next to c' and that the other community holds true the sentence 'there are 9 objects next to c'. In addition, both sentences turn out to be true, given that there are actually 9 objects next to c. So, point (ii*) is taken care of.

Point (i) of the definition requires the two languages to be non-intertranslatable. This is probably the requirement which needs a more elaborate argument. How can one show that two languages are non-intertranslatable? To do that, I suggest to characterize the two lan-

[16] For expository purposes I assume their languages to be very poorly expressive. There is no problem in doing so. Definition (INC-DS*) does not restrict the choice of supposed incommensurable languages to very expressive one, although it can be debated whether a definition of incommensurability should do so.

[17] This should not come as a too big surprise, because the concept corresponding to the number zero, after all, is a rather abstract one and it is possible that communities of speakers lack it, although they can master the other numbers.

guages in a formal fashion, so that failure of intertranslatability can be argued rigorously.[18]

The language which can only express the exact number of objects next to c will be denoted by L_n, while $L_\exists$ will denote the language which can only express whether there are, or there aren't , objects next to c. The syntax and semantics of L_n and $L_\exists$ should be defined first, and subsequently their non intertranslatability can be argued.

Definition **1** (Syntax)**.** *In terms of vocabulary, languages $L_\exists$ and L_n share the constant 'c,' the variable 'x,' the two-place predicate 'next-to,' and connectives '$\neg$' and '$\vee$.' They differ in their vocabulary because $L_\exists$ contains the quantifier '$\exists$' (there are. . .), whereas L_n contains the quantifier '$\exists_n$' (there are exactly n. . .).*

The well-formed formulas of L_n and $L_\exists$ can be atomic or complex. $L_\exists$ contains only one atomic formula, namely '$\exists x : \textit{next-to}(x, c)$.' Instead, there is an n number of atomic formulas in L_n and they are only of the shape '$\exists_n x : \textit{next-to}(x, c)$,' with n any natural number such that $n > 0$. Complex formulas of $L_\exists$ and L_n are built recursively using the connectives $\neg$ and $\vee$.

Definition **2** (Semantics)**.** *Let $|\ldots|_n$ and $|\ldots|_\exists$ be interpretation functions from elements of the vocabulary of L_n and $L_\exists$ to objects of an infinite domain D:*

- $|c|_n = |c|_\exists = \overline{c}$ *for a given* $\overline{c} \in D$;
- $|\textit{next-to}|_n = |\textit{next-to}|_\exists = \{(d, \overline{c})\colon d \in N \subseteq D \& d \neq \overline{c}\}$.

Next, truth-conditions for formulas in $L_\exists$ and L_n can be defined as follows,where # expresses the cardinality of a set:

- $\exists x : \textit{next-to}(x, c)$ *is true iff* $\#\{d \in D\colon (d, \overline{c}) \in |\textit{next-to}|_\exists\} \geq 1$;
- $\exists_n x : \textit{next-to}(x, c)$ *is true iff* $\#\{d \in D\colon (d, \overline{c}) \in |\textit{next-to}|_n\} = n$, *for any* n;
- *the recursive clauses for $\vee$ and $\neg$ are standard.*

Claim **1.** *Languages L_n and $L_\exists$ are totally non-intertranslatable, if $\top$ and $\perp$ are omitted.*

Recall that two languages are totally non-intertranslatable iff for *all* sentences in L, no meaning-equivalent sentence in L' can be found, and viceversa. But what does it take for two sentences or formulas to be meaning-equivalent? Following Davidson, I will assume that two

[18] One could object that the formal renderings of the two languages talked by the communities is not correct. The reader may judge by herself whether or not the formal renderings I am going to propose have misrepresented the languages of the two communities.

sentences or formulas have the same meaning if they have the same truth-conditions.[19]

Preliminaries aside, the argument for claim 1 can now be given. I shall here provide only the semi-formal and intuitive ideas on the assumption that formal details can be worked out. Two sub-claims should be established: (a) for any $\varphi_\exists$, if $\varphi_\exists \in L_\exists$, then $\varphi_\exists$ cannot be translated into any $\varphi_n \in L_n$, provided $\varphi_\exists$ is not $\top$ or $\bot$; and (b) for any φ_n, if $\varphi_n \in L_n$, then φ_n cannot be translated into any $\varphi_\exists \in L_\exists$, provided φ_n is not $\top$ or $\bot$.

First, the argument for (a): Consider the formula '$\exists x : \mathit{next\text{-}to}(x, c)$' in $L_\exists$. The sentence would correspond to an infinite disjunction in L_n such as '$\exists_1 x : \mathit{next\text{-}to}(x, c) \vee \exists_2 x : \mathit{next\text{-}to}(x, c) \vee \ldots$.' However, infinite disjunctions are not allowed in L_n. Likewise, '$\neg\exists x : \mathit{next\text{-}to}(x, c)$' cannot be translated into L_n, since we assumed $n > 0$. Taking disjunctive formulas in $L_\exists$ would not change much, for these will be either equivalent to some atomic or negated atomic formulas in $L_\exists$, or they will be equivalent to $\top$ or $\bot$.

Next, the argument for (b): Consider the formula '$\exists_n x : \mathit{next\text{-}to}(x, c)$' in L_n. Any sentence in $L_\exists$ such as '$\exists x : \mathit{next\text{-}to}(x, c)$' cannot work as a good translation. For instance consider the case there are $n + 1$ objects next to w; then, '$\exists_n x : \mathit{next\text{-}to}(x, c)$' would be false but '$\exists x : \mathit{next\text{-}to}(x, c)$' would be true. Likewise, a negated formula such as '$\neg\exists x : \mathit{next\text{-}to}(x, c)$' would not work either, as the latter is true iff 0 objects are next to c. Again, taking disjunctive formulas in L_n would not change much, for these will be either equivalent to to $\top$ or $\bot$, or they could be shown to be untranslatable by the same argument used for the atomic formulas. This establishes the total non-intertranslatability between L_n and $L_\exists$.

6.5.2 Against the second reading argument

The refutation of Davidson's argument under the first reading, and thus of claim (1), can hardly be resisted given its formal fashion. However, one may well think that the first reading is not the correct reconstruc-

[19] This is one of the points in the discussion about semantical incommensurability when an appeal to a theory of meaning is crucial. Davidson himself can be seen as advocating a truth-conditional account of meaning, for which formulas in L_n and $L_\exists$ are meaning-equivalent whenever one formula is true iff the other is true. I do not want to defend this assumption in full, but pointing out its plausibility should suffice. With regard to L_n and $L_\exists$, the assumption that meaning of formulas boils down to truth-conditions is reasonable on the ground that that L_n and $L_\exists$ are descriptive languages composed of statements about the world. They are not composed of pieces of discourse such as commands or imperatives which are typically troublesome for a truth-conditional account of meaning.

tion of what Davidson had in mind. So let me now move to the second reading and the refutation of claim (2).

Consider, once again, languages $L_\exists$ and L_n and suppose condition (C) is satisfied, *i.e.*, the incommensurability claim takes place in, say, language $L_\exists$. I contend that speakers in $L_\exists$ can coherently claim that their language is incommensurable to L_n. Suppose that language L_n is totally alien to the speakers of $L_\exists$, or that speakers of $L_\exists$ have no idea what the speakers of L_n mean when they communicate. Further, suppose that speakers of $L_\exists$ are trying to find out a translation of L_n into their own language. This situation is that of a radical interpretation.

Given the radical interpretation scenario, I shall argue for two claims. My first claim is that speakers of $L_\exists$ are able to arrive at the conclusion that L_n is *untranslatable* into their own language. My second claim is that speakers of $L_\exists$ are able to arrive at the conclusion that sentences of L_n are true. If both claims are warranted, then speakers of $L_\exists$ can arrive at the conclusion that L_n and $L_\exists$ are incommensurable in the sense of (INC-D*). To anticipate, I must say that speakers of $L_\exists$ can only arrive at the conclusion that L_n and $L_\exists$ are *partially* incommensurable, because they can only arrive at the conclusion that L_n and $L_\exists$ are *partially* non-intertranslatable. But this is good enough to undermine Davidson's argument.

I will start by arguing for my first claim. Davidson grants that in a situation of radical interpretation we can make two assumptions. First: (A1) no matter which language we speak, we can always tell whether someone is holding a sentence true or not. Second: (A2) we can question someone by asking whether she holds a sentence true (provided the sentence belongs to the language she speaks), and we can also understand her reaction of assent or dissent. This is crucial for the process of radical interpretation to get started.[20]

Suppose now that speakers of $L_\exists$ hear speakers of L_n utter sentence φ_n. Then, speakers of $L_\exists$ will formulate an hypothesis about the meaning of φ_n by associating it with some sentence $\varphi_\exists$ in their own language $L_\exists$. The hypothesis may have the following form:

$$\varphi_n \text{ means the same as } \varphi_\exists.$$

With Davidson we may assume that speakers of $L_\exists$ have a truth-conditional theory of the sameness of meaning, hence the hypothesis will look like the following:

(H) $$`\varphi_n\text{' is true iff } \varphi_\exists$$

[20]Davidson writes: "Suppose , then, that the evidence available is just that speakers of the language to be interpreted hold various sentences to be true at certain times and under specific circumstances," (Davidson, 1978, p. 135).

Speakers of $L_\exists$ can test (H) by questioning speakers of L_n. How does this work? First, speakers of $L_\exists$ have to decide whether $\varphi_\exists$ is true or not, and they can do so by knowing what $\varphi_\exists$ means and depending on the state of the world given a time and a space location. Next, they will ask whether speakers of L_n hold φ_n true or not. If they discover that there is a situation in which $\varphi_\exists$ is true but φ_n is not held true by speakers of L_n, that would show that $\varphi_\exists$ is not a good translation of φ_n.

Let us look at an example. Speakers of $L_\exists$ hear the sentence '$\exists_9 x$: *next-to*(x, c)' from speakers of L_n. Now, they will formulate an hypothesis and try to test it:

(H9) '$\exists_9 x$: *next-to*(x, c)' is true iff $\exists x$: *next-to*(x, c).

Speakers of $L_\exists$ will then consult speakers of L_n under the assumption that such a consultation is possible between speakers of L_n and $L_\exists$. Suppose that the situation is such that there are 9 objects next to c. Speakers of $L_\exists$ hold the sentence '$\exists x$: *next-to*(x, c)' true, and by (H9) they expect speakers of L_n to hold '$\exists_9 x$: *next-to*(x, c)' true. Speakers of L_n will obviously assent. This would be a first piece of evidence that '$\exists x$: *next-to*(x, c)' is a translation for '$\exists_9 x$: *next-to*(x, c)'. Now, one piece of evidence is not enough, and so speakers of $L_\exists$ will test (H9) another time, and suppose that this time the situation is such that there are 8 objects next to c instead of 9. Speakers of $L_\exists$ would still hold '$\exists x$: *next-to*(x, c)' true and thus by (H9) they would expect speakers of L_n to hold '$\exists x$: *next-to*(x, c)' true. But this time speakers of L_n will show their dissent.

From the evidence gathered, speakers $L_\exists$ will be able to conclude that the translation hypothesis (H9) is false and they will try another hypothesis. However, in general we know by claim 1 that any translation hypothesis will fail, *i.e.*, any attempt to translate '$\exists_9 x$: *next-to*(x, c)' into a sentence of $L_\exists$ will fail. So, suppose that speakers of $L_\exists$ can go through all the sentences in their language;[21] then, they will conclude that '$\exists_9 x$: *next-to*(x, c)' has no translation in their language. This will show that speakers of $L_\exists$ can arrive at the conclusion that their language is *partially* incommensurable to L_n.[22]

[21] After all, there are not very many sentences: $\exists x$: *next-to*(x, c), $\neg\exists x$: *next-to*(x, c), and the rest are tautologies and contradictions which are always translatable under a truth-conditional account of meaning, and thus are to be disregarded.

[22] A stronger conclusion could be reached, though. Speakers of $L_\exists$ wil reach the same untranslatability conclusion for any sentence they will possibly hear from speakers of L_n. So, speakers of $L_\exists$ will reach the conclusion that, *as far as they know*, any sentence in L_n cannot be translated by any sentence in $L_\exists$. Note that

This reasoning establishes my first claim, namely that speakers of $L_\exists$ can conclude that their language is not translatable (at least partially) into L_n. Yet some may object that the reasoning utilized by speakers of $L_\exists$ to reach the conclusion that $L_\exists$ and L_n are partially non-intertranslatable exceeds the expressive power of $L_\exists$ itself, so the conclusion that L_n and $L_\exists$ are untranslatable cannot be formulated into $L_\exists$ itself, nor can the reasoning leading to such a conclusion. True enough. Let us then extends $L_\exists$ in such a way that the untranslatability claim can be formulated. So, let $L_\exists^+$ be an extension of $L_\exists$ by adding: the predicate 'is true', the bi-conditional 'iff', and names to denote sentences heard from speakers of L_n. This will allow speakers of $L_\exists^+$ to formulate type (H) hypotheses. Moreover, speakers of $L_\exists^+$ will need a rule (R) of the form: *If any type (H) hypothesis of translation for a sentence φ_n in L_n encounters the dissent of speakers of L_n, then φ_n is not translatable into any sentence in $L_\exists^+$*. So, once $L_\exists$ is extended this way, it seems that we would have a scenario in which (C) is satisfied, and in which speakers of $L_\exists^+$ can coherently claim that their language is not translatable into L_n, at least partially.[23]

The second claim to be established is that speakers of $L_\exists$ can assert that languages L_n and $L_\exists$ (or better, subsets of them) are true. This is the point where Davidson's argument will start to apply. Recall: Truth is based on translatability according to Tarski's (T) convention, so any sentence that is claimed to be true by the speakers of a language will be translatable into that language. Hence, if speakers of $L_\exists$ claims that a sentence φ_n of L_n is true, that sentence will be translatable into some sentence in $L_\exists$. How can this argument be resisted? My strategy would be to undercut the very premise of the argument, namely that truth is based on translatability. I will argue that Davidson cannot consistently hold that truth is based on translatability, and at the same time hold other theses that are essential to his philosophy of language. The crucial point consists in assuming another Davidsonian thesis, *i.e.*, (A3) truth

this conclusion is *weaker than total* non-intertranstabality since speakers of $L_\exists$ can never be sure that the sentences their heard from speakers of L_n are alla sentences that are in L_n.

[23] There is a residual objection: The non-intertranslatablity claim was established for $L_\exists$ by using the additional expressive power of $L_\exists^+$. But what we need is to establish the non-intertranslatability claim for $L_\exists^+$ by using the expressive power of $L_\exists^+$ itself. However, I think that although $L_\exists^+$ has additional expressive power over $L_\exists$, this is of no use for finding sentences in $L_\exists^+$ that are meaning-equivalent to sentences in L_n. Observe: the sentences that are in $L_\exists^+$ but not in $L_\exists$ are type (H) hypotheses and the rule (R), but these sentences cannot serve as a translation for sentences in L_n.

precedes meaning,[24] and see that it conflicts with the claim that truth is based on translatability.[25]

It is correct to say that given a Tarski's bi-conditional, truth is based on translatability. However, Davidson has inverted the relation between translatability and truth that is encoded in Tarski's conventions (T), by inverting the relation between meaning and truth according to (A3). In Davidson, Tarski's bi-conditionals are used to deliver a theory of meaning based on truth, and not a theory of truth based on meaning. Hence, if truth precedes meaning, truth precedes translatability. My conclusion is, then, that Davidson claim that truth is based on translatability is correct if applied to Tarski (T) conventions as such, but it is not correct if applied to the way in which Davidson uses Tarski (T) conventions to deliver a theory of meaning, and not a theory of truth.[26]

More precisely, the relation between truth, meaning and translatability should be as follows. In the process of radical interpretation, one first recognizes that a sentence is held true, and also registers all cases in which that sentence is held true. Next, by the Davisonian principle of charity, one takes a sentence that is held true as simply a true sentence. Finally, one determines what the meaning of the sentence is on the basis of the circumstances in which it is (held) true. Clearly, in the process of radical interpretation one arrives at determining the meaning of a sentence that belongs to an alien language by finding a meaning-equivalent sentence into her own familiar language, or by finding a translation of the alien sentence into a sentence of her own familiar language. But notice that, in order for the whole process to get started, one has to have an independent understanding of what it takes for a sentence to be (held) true; and Davidson grants that holding a sentence true is a primitive attitude which we can all understand across languages. Under a Davidsonian standpoint, thus, there seems to be no problem in assuming that speakers of $L_\exists$ can understand and claim that some sentences of L_n are true without yet possessing a translation of these sentences into their own language.

[24]See in particular Davidson (1967).

[25]Assumption (A3) was also granted in the process of radical interpretation between $L_\exists$ and L_n: The attitude of holding a sentence of L_n true is recognized by the speakers of $L_\exists$ before the meaning of that sentence is reconstructed.

[26]These ideas are in line with Davidson (1973). Davidson himself writes: "In Tarski's work, T-sentences are taken to be true because the right branch of the bi-conditional is assumed to be a translation of the sentence truth conditions for which are being given ... What I propose is to reserve the direction of explanation: assuming translation, Tarski was able to define truth; the present idea is take truth as basic and to extract an account of translation or interpretation." (Davidson, 1978, p. 134).

To summarize, I have shown that speakers of $L_{\exists}^{+}$ can claim that their language is untranslatable to L_n. They can also claim that sentences of both $L_{\exists}^{+}$ and L_n are true by relying on the primitive notion of holding sentences true. As a result, speakers of $L_{\exists}^{+}$ can claim that their language is incommensurable to L_n, whence claim (2) fails.

6.5.3 Against the third reading argument

It remains to be established that claim (3) fails. More modestly, I will argue that Davidsons does not provide us with any evidence against the denial of claim (3). For claim (3) to fail, speakers of English should be able to claim that another language, call it Alien, is untranslatable into English, and moreover that sentences of English and Alien are true. Suppose speakers of English are confronted with speakers of Alien. By the same argument given for the second reading, from the attitude of holding true and by applying the principle of charity, speakers of English will be able to claim that sentences of Alien are true, although they do not know what they mean. It is significantly more difficult to show that speakers of English can claim that Alien and English are non-intertranslatable. Here I can only point out a parallelism. In the imaginary radical translation scenario, speakers of $L_{\exists}$ encountered speakers of L_n, and by testing several translation hypotheses (H), they could conclude that $L_{\exists}$ and L_n were non-intertranslatable (at least partially). Similarly, one can imagine speakers of English encountering speakers of Alien and realizing that Alien is untranslatable into English, in the same way in which speakers of $L_{\exists}$ realized that L_n is not translatable into $L_{\exists}$. This parallelism gives us the conclusion:

(E1) *Possibly*, members of our community of speakers can coherently claim that our language is not-translatable into another.

But it does give use the conclusion:

(E2) Members of our community of speakers can coherently claim that our language is not-translatable into another.

From (E1) and the fact argued before that speakers of English can claim that sentences of both English and Alien are true, this conclusion follows:

(E3) *Possibly*, members of our community of speakers can coherently claim that our language is incommensurable to another.

One can see that (E3) is not the outright denial of (3), unless (3) is read as a necessitated claim. As a result, my conclusion is that Davidson does not offer any evidence against the possibility of claim (3) failing, yet it is fair to say that I did not offer any conclusive evidence for the actual failure of (3).

6.6 Conclusion

In this paper I have reconstructed Davidson's definition of incommensurability in the form of (INC-D*). I have given three readings of his argument against incommensurability, and for each reading I have given counter-arguments showing that the argument cannot go through. With regard to the first reading, I have constructed two languages that are incommensurable according to (INC-D*). With regard to the second reading, I have shown that Davidson's argument cannot go through by assumptions (A1), (A2) and (A3). My conclusion has been weaker in the case of the second reading than in the case of the first reading. Under the latter, I could show that two given languages are incommensurable, whereby refuting claim (1). Under the former, I could show that speakers of one language can claim that their language is partially (but not totally) incommensurable to another language, whereby refuting claim (2), but only if incommensurability is taken to be partial incommensurability. In the case of the third reading, I could only offer a conjecture that claim (3) fails, in the sense that there is no available evidence to deny the failure of (3).

To conclude, I would like to address two general lines of reply that are open to Davidson. One is that languages $L_\exists$ and L_n are not natural languages, and that Davidson's argument was only about natural languages. This is a pressing worry. However, it is unclear to me how any discussion about the incommensurability between natural languages can be made precise in the first place. In particular, arguing that a natural language is non-intertranslatable into another natural language will require us to have a full account of meaning in natural languages. But this account is lacking for now. Thus, either we can circumvent the problem (but how?), or adopting toy-languages such as $L_\exists$ and L_n seems—for the time being—unavoidable, or at least convenient.

The second line of response is that languages $L_\exists$ and L_n are very poorly expressive, and that if more expressive languages were considered the incommensurability result would not be yielded. This is an interesting conjecture. For instance, it would be interesting to be able to show that, if two languages reach a certain threshold of expressive power, they will be *eo ipso* intertranslatable and thus commensurable. However, until this threshold of expressive power is spelled out precisely, I do not see any reason why very poorly expressive languages should not be considered.

References

Davidson, Donald. 1967. Truth and meaning. *Synthese* 17:304–323.

Davidson, Donald. 1973. Radical interpretation. *Dialectica* 27(3-4):313–328.

Davidson, Donald. 1974. On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association* 47:5–20.

Davidson, Donald. 1978. *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

Hacker, P. M. S. 1996. On davidson's idea of conceptual scheme. *The Philosophical Quarterly* 46(184):289–307.

Sankey, Howard. 1993. Kuhn's changing concept of incommensurability. *British Journal for the Philosophy of Science* 44(4):775–791.